Impact of Windowing on the Detection Probability of Weak Sine Waves Affected by Frequency Fluctuation

Diego Bellan

Abstract—This work investigates the impact of windowing on the detection probability of weak sine waves buried in noise and characterized by a frequency uniformly distributed in a given frequency interval. Detection is performed in the frequency domain through the discrete Fourier transform (DFT). The mean value of the sine wave frequency can be located in any position on the frequency axis, i.e., it is not required a frequency mean value corresponding to a DFT bin. The work extends previous results obtained under the simple worst-case assumption of a sine wave frequency located in the middle point between two adjacent DFT bins. An interesting application related to such extension is given by the detection of interharmonics in power systems. Analytical expressions for the detection probability as function of the noise level and the selected window are derived in closed form. Validation is performed through numerical simulation of the whole measurement process.

Keywords—Detection probability, discrete Fourier transform, frequency fluctuation, noise, weak sine waves.

I. INTRODUCTION

Detection of weak sine waves buried in additive noise arises in many fields of electrical engineering, including aerospace applications, diagnostics of electrical machines, and measurement of radiated electromagnetic emissions (e.g., see [1]-[3]). Although it is well known that parametric methods provide better performances in sine wave detection (e.g., see [4]), in many applications the nonparametric method based on the discrete Fourier transform (DFT) is still the most common approach due to its simplicity and since it does not require a preliminary selection of the order of the spectral model. As far as the DFT approach is considered, windowing against spectral leakage must be taken into account when sine wave detection is investigated [1], [4]-[6]. In fact, windowing greatly affects sine wave detection for several reasons. First, it is well known that different windows result in different noise level in each DFT frequency bin. Second, spectral leakage due to asynchronous sampling results in interference between spectral lines. Third, frequency resolution of the windowed DFT could be lower than the frequency spacing of adjacent spectral lines. The three points mentioned above have been already investigated in the related literature, therefore in this paper it is assumed that: 1) The noise behavior of a window is properly taken into account by the corresponding equivalent noise bandwidth (ENBW); 2) Side-lobes level of each considered window is such that spectral leakage can be neglected; 3) Frequency resolution of DFT is such that the sine wave to be detected can be distinguished from the adjacent spectral lines. A further point related to windowing, however, must be taken into account when sine wave detection is considered. Under asynchronous sampling conditions, the amplitude of the sine wave to be detected is weighted by the main lobe of the window in the frequency domain because the sine wave frequency deviates from the center of the window main lobe [5], [7]-[8]. Such weighting process results in attenuation of the sinewave amplitude, leading to lower detection probability. In the worst case, i.e., when the sine wave frequency is placed in the middle point between two adjacent frequency bins, the maximum amplitude attenuation is given by the window scallop loss (SL) [5]. This phenomenon is well known in signal processing theory, but its impact on the detection probability of a sine wave was not thoroughly investigated in the existing literature. In fact, in the literature only the worst case (i.e., the maximum attenuation SL) was considered [1], whereas a proper investigation requires the analytical description of the impact of the window main-lobe continuous behavior in the frequency domain. To this aim, in this paper it is assumed that the frequency of the sine wave to be detected is randomly distributed within a given frequency interval with uniform distribution (i.e., frequency fluctuation about a frequency mean value is assumed). By taking into account the frequency behavior of the window main lobe introducing amplitude attenuation, the probabilistic description of the spectral line amplitude corresponding to the sine wave is derived in closed form. Finally, the detection probability is derived by taking into account the noise behavior of the selected window. The results presented in the paper extend the results presented in [7]-[8] where the special case of a harmonic sine wave was considered, i.e., where the frequency mean value of the fluctuating sinewave was equal to the center of a frequency bin. On the contrary, in this paper the mean value of the sine wave frequency can be located in any position. An important application of such assumption is related to detection of interharmonics in power systems [9]-[10].

The paper is organized as follows. In Section II the basic results concerning detection and false alarm probabilities in

D. Bellan is with the Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milan, Italy (e-mail: diego.bellan@polimi.it).
the frequency domain already available in the literature are recalled. In Section III the approximate analytical model for the frequency behavior of weighting windows is introduced, and the probabilistic description of frequency fluctuation effects is analytically derived. In Section IV the approach based on detection and false alarm probabilities of a noisy sine wave component is extended taking into account sine wave frequency fluctuation. In Section V the mean value and the variance of the spectral line amplitude are derived. Analytical results are validated through numerical simulations in Section VI.

II. BACKGROUND: PROBABILITY DISTRIBUTION OF DFT SPECTRAL LINES AMPLITUDE

Let us consider a distorted voltage or current waveform \( u(t) \) consisting in the sum of \( N \) sine waves and zero-mean white noise \( n(t) \):

\[
  u(t) = \sum_{h=1}^{N} A_h \cos(2\pi f_h t + \phi_h) + n(t) \quad (1)
\]

The waveform \( u(t) \) is sampled with sampling period \( T_s \), i.e., with sampling frequency \( f_s = 1/T_s \). The time-to-frequency transformation of the sampled waveform is performed by evaluating the \( N_S \) samples discrete Fourier transform (DFT) [4]-[6]:

\[
  U_m = \frac{2}{N_S \text{NPSG}} \sum_{k=0}^{N_S-1} u(kT_s) w(kT_s) \exp \left( -j \frac{2\pi mk}{N_S} \right) \quad (2)
\]

where \( w(t) \) is a window function used against spectral leakage, and NPSG is the normalized peak signal gain of the selected window \( w(t) \). The frequency resolution of the DFT in (2) is \( \Delta f = f_s/N_S \). The factor \( 2/(N_S \text{NPSG}) \) in (2) is such that the estimate of each sine wave amplitude in (1) is given by the coefficient magnitude \( |U_m| \) corresponding to the discrete frequency \( m \Delta f \) closest to the considered sine wave frequency.

Additive noise \( n(t) \) results in a random behavior of the DFT coefficients \( U_m \). The real and the imaginary parts of each \( U_m \) can be approximated as a Gaussian random variable (RV) with unbiased mean values (i.e., the deterministic noise-free values), and variance [6]-[10], [11]-[17]

\[
  \sigma^2 = \text{ENBW} \frac{2}{N_S} \sigma_n^2 \quad (3)
\]

where ENBW is the equivalent noise bandwidth of the selected window \( w(t) \), and \( \sigma_n^2 \) is the variance of the input noise \( n(t) \). As a result, the only-noise spectral lines (i.e., with zero deterministic component) have a Rayleigh probability density function (PDF) [18]-[20]:

\[
  g_{|U|}(|U|) = \frac{|U|}{\sigma^2} \exp \left( -\frac{|U|^2}{2\sigma^2} \right) \quad (4)
\]

and cumulative distribution function (CDF):

\[
  G_{|U|}(|U|) = 1 - \exp \left( -\frac{|U|^2}{2\sigma^2} \right) \quad (5)
\]

Notice that in (4) and (5) the frequency subscripts \( m \) was dropped since the results are independent of \( m \).

On the other hand, the signal spectral lines (i.e., with non-zero deterministic component) have Rician PDF [7]-[8], [11]-[20]:

\[
  f_{|U|}(|U|) = \frac{|U|}{\sigma^2} \exp \left( -\frac{|U|^2 + C^2}{2\sigma^2} \right) I_0 \left( \frac{|U| C}{\sigma^2} \right) \quad (6)
\]

where \( C \) is the related sine wave amplitude (i.e., \( A \) with proper index) weighted by the spectrum of the window \( w(t) \), and \( I_0 \) is the modified Bessel function of the first kind. If the considered sine wave frequency equals an integer multiple of \( \Delta f \) than the weight introduced by the window spectrum is equal to one, otherwise the weight is less than one. The CDF is given by

\[
  F_{|U|}(|U|) = 1 - Q_1 \left( \frac{C \sigma}{\sigma^2} \right) \quad (7)
\]

where \( Q_1 \) is the Marcum Q function.

By defining a threshold level \( \alpha \), the false alarm probability is defined as the probability that an only-noise spectral line is larger than \( \alpha \). Thus, from (5) the false alarm probability is given by

\[
  P_f(\alpha) = \exp \left( -\frac{\alpha^2}{2\sigma^2} \right) \quad (8)
\]

The detection probability is defined as the probability that a signal spectral line is greater than the threshold. Thus, from (7) the detection probability is given by

\[
  P_d(\alpha) = Q_1 \left( \frac{C \alpha}{\sigma^2} \right) \quad (9)
\]

By solving (8) with respect to \( \alpha \) and substituting into (9) we obtain

\[
  P_d(P_f(\alpha)) = Q_1 \left( \frac{C}{\sqrt{2\log(P_f(\alpha))}} \right) \quad (10)
\]

Therefore, for a given signal-to-noise ratio \( C/\sigma \), eq. (10) provides the detection probability as a function of the accepted false alarm probability.

In the next Sections the PDF and the CDF of spectral lines affected by frequency fluctuation will be derived instead of (4) and (5), respectively. Therefore, the statistical moments (e.g., mean value and variance) and the detection probability similar to (10) will be derived taking into account the random distribution of sine wave frequency and the weighting window used against spectral leakage.

III. PROBABILISTIC MODELING OF SINE WAVE FREQUENCY FLUCTUATION

It is well known that when a DFT is evaluated, a time-domain sine wave component results in a spectral line whose amplitude is weighted by the main lobe of the spectrum of the window function \( w(t) \) used against spectral leakage. Therefore,
by assuming negligible spectral leakage, a sine wave amplitude is weighted by a frequency-domain function equal to one at the DFT frequency bins (i.e., integer multiples of \( \Delta f \)), and equal to the scallop loss (SL) of the selected window at \( (m \pm 1/2)\Delta f \). The behavior of such function within the above mentioned frequency range is specific to the selected window. In order to obtain approximate and simple results, however, the frequency behavior of the window function can be analytically represented by the parabolic behavior (see Fig. 1) [7]-[8]

\[
y \equiv 1 - \frac{4(1-SL)}{\Delta f^2} f^2 = 1 - 4(1-SL)x^2
\]

(11)

where the origin was assumed at a frequency bin \( m\Delta f \), and \( x = f/\Delta f \) is the normalized frequency. Therefore, after DFT the amplitude of a spectral line can be written as

\[
A(x) = y(x) \cdot A_0
\]

(12)

where \( A_0 \) is the actual amplitude of the sine wave component (i.e., one of the coefficients \( A_b \) in (1)).

In this paper, the sine wave frequency is treated as a random variable (RV) uniformly distributed within a given (normalized) frequency range \( \Delta x = x_2 - x_1 < 1 \), centered around the mean normalized frequency \( x_0 \) (i.e., the normalized deviation with respect to a frequency bin). In the following, the derivations are performed by assuming \( 0 \leq x_0 \leq 1/2 \) since \( y(x) \) is an even function and therefore the same results can be obtained for negative \( x_0 \). Thus, the spectral line amplitude (12) can be regarded as a function of the RV \( x \), and the statistical properties of the RV \( y \) can be readily obtained from Fig. 1. In particular, the CDF of \( y \) can be obtained by evaluating the frequency interval where \( y \leq \bar{y} \) for \( SL \leq \bar{y} \leq 1 \). To this aim, four different cases should be investigated:

1) \( x_1 \geq 0, x_2 \leq 1/2 \). By taking into account that

\[
\bar{x} = \frac{1-y}{\sqrt{4(1-SL)}}
\]

(13)

the CDF of \( y \) can be readily obtained as

\[
F_y(y) = \frac{x_2-x}{\Delta x} = \frac{x_0}{\Delta x} + \frac{1}{2} - \frac{1}{2\Delta x} \sqrt{\frac{1-y}{4(1-SL)}}, \quad y_2 \leq y \leq y_1,
\]

(14)

where \( y_1 = 1 - 4(1-SL)x_1^2 \) and \( y_2 = 1 - 4(1-SL)x_2^2 \). The PDF of \( y \) is obtained as the first derivative of (14):

\[
f_y(y) = \frac{dy}{dy} = \frac{1}{4\Delta x \sqrt{1-SL} \sqrt{1-y}} = p_0(y)
\]

(15)

2) \( x_1 < 0, x_2 \leq 1/2 \). In this case it can be shown that the CDF of \( y \) is given by:

\[
F_y(y) = \left\{ \begin{array}{ll}
\frac{x_0}{\Delta x} - \frac{1}{2} - \frac{1}{2\Delta x} \sqrt{\frac{1-y}{1-SL}}, & y_2 \leq y \leq y_1 \\
1 - \frac{1}{\Delta x} \sqrt{\frac{1-y}{1-SL}}, & y_1 \leq y \leq 1
\end{array} \right.
\]

(16)

and the related PDF:

\[
f_y(y) = \left\{ \begin{array}{ll}
p_0(y), & y_2 \leq y \leq y_1 \\
2p_0(y), & y_1 \leq y \leq 1
\end{array} \right.
\]

(17)

3) \( x_1 > 0, x_2 > 1/2 \) (i.e., the case shown in Fig. 1). The CDF and PDF are given by:

\[
F_y(y) = \left\{ \begin{array}{ll}
\frac{1}{2\Delta x} - \frac{1}{2\Delta x} \sqrt{\frac{1-y}{1-SL}}, & SL \leq y \leq y_2 \\
1 - \frac{1}{\Delta x} \sqrt{\frac{1-y}{1-SL}}, & y_1 \leq y \leq 1
\end{array} \right.
\]

(18)

\[
f_y(y) = \left\{ \begin{array}{ll}
2p_0(y), & SL \leq y \leq y_2 \\
p_0(y), & y_1 \leq y \leq y_1
\end{array} \right.
\]

(19)

4) \( x_1 < 0, x_2 > 1/2 \). The CDF and PDF are given by:

\[
F_y(y) = \left\{ \begin{array}{ll}
\frac{1}{2\Delta x} - \frac{1}{2\Delta x} \sqrt{\frac{1-y}{1-SL}}, & SL \leq y \leq y_2 \\
1 - \frac{1}{\Delta x} \sqrt{\frac{1-y}{1-SL}}, & y_1 \leq y \leq 1
\end{array} \right.
\]

(20)

\[
f_y(y) = \left\{ \begin{array}{ll}
2p_0(y), & SL \leq y \leq y_2 \\
p_0(y), & y_1 \leq y \leq y_1
\end{array} \right.
\]

(21)

Finally notice that from (12) the CDF and PDF of the RV \( A \) can be obtained as:

\[
F_A(A) = F_y \left( \frac{A}{A_0} \right)
\]

(22)
\[ f_A(A) = \frac{1}{A_0} f_y \left( \frac{A}{A_0} \right) \quad (23) \]

IV. DETECTION PROBABILITY OF SINE WAVE

The measured amplitude of a spectral line is affected by both additive noise and frequency fluctuation. Therefore, the detection probability for a given threshold level \( \alpha \) must be obtained from the total probability theorem by combining (9) (representing only the noise contribution for a given sine wave amplitude \( A \)) and (23) (representing the frequency fluctuation contribution) [18]-[20]:

\[ P_d(\alpha) = \int_{SL-A_0}^{A_0} f_d(\alpha | A) f_A(A) dA = \]

\[ = \int_{SL-A_0}^{A_0} Q_1 \left( \frac{A}{\sigma} \right) f_A(A) dA \quad (24) \]

and by taking into account (12) we obtain:

\[ P_d(\alpha) = \int_{SL}^{1} Q_1 \left( SNR \cdot \gamma \frac{\sigma}{\sigma} \right) f_\gamma(y) dy \quad (25) \]

where

\[ SNR = \frac{A_0}{\sigma} \quad (26) \]

is the signal-to-noise ratio.

Finally, by taking into account (8), the detection probability can be expressed as a function of the false alarm probability as

\[ P_d(P_{fa}) = \int_{SL}^{1} Q_1 \left( SNR \cdot \gamma \sqrt{-2\log(P_{fa})} \right) f_\gamma(y) dy \quad (27) \]

In (27) the PDF \( f_\gamma(y) \) is given by (15), (17), (19), or (21) according to the frequency range \( (x_1, x_2) \).

V. MEAN VALUE AND VARIANCE OF MEASURED SINE WAVE AMPLITUDE

By denoting as \( M = |U| \) the measured amplitude of a spectral line affected by both additive noise and frequency fluctuation, the related PDF can be obtained by resorting to the total probability theorem and by using (6) and (23):

\[ f_M(M) = \int_{SL-A_0}^{A_0} f_M(M | A) f_A(A) dA = \]

\[ = \int_{SL-A_0}^{A_0} \frac{M}{a^2} \exp \left( -\frac{M^2 + A^2}{2a^2} \right) I_0 \left( \frac{MA}{a} \right) f_A(A) dA \quad (28) \]

By taking into account (12) and by defining the normalized amplitude of the measured sine wave:

\[ z = \frac{M}{\sigma} \quad (29) \]

we obtain:

\[ f_z(z) = \int_{SL}^{1} z \cdot \exp \left( -\frac{z^2 + SNR \cdot \gamma^2}{2} \right) I_0 \left( \frac{SNR \cdot \gamma y}{\sigma} \right) f_\gamma(y) dy \quad (30) \]

The mean value of \( M \) can be evaluated as:

\[ \mu_M = \sigma \mu_z = \sigma \int_0^\infty z f_z(z) dz \quad (31) \]

and the variance:

\[ \sigma_M^2 = \sigma^2 \mu_z^2 = \sigma^2 \int_0^\infty (z - \mu_z)^2 f_z(z) dz \quad (32) \]

VI. NUMERICAL VALIDATION

The analytical results derived in Sections IV and V were validated by resorting to numerical simulation of the whole measurement process. According to (1), a waveform consisting of three harmonic components was selected such that \( f_1 = 50 \text{ Hz}, \ f_2 = 150 \text{ Hz}, \ f_3 = 250 \text{ Hz}. \) The amplitudes were selected as \( A_1 = 10, A_2 = 2, A_3 = 1, \) whereas the phase angles were selected at random. One sine wave component \( A_4 \) was also added, with uniformly distributed random frequency. Several values for the average frequency were selected within the range \((180 \text{ Hz}, 182.5 \text{ Hz}), \) i.e., in normalized units within the range \((36, 36.5). \) Different values were also selected for the frequency interval \( \Delta x = x_2 - x_1, \) provided that \( \Delta x < 1. \) The sine wave amplitude \( A_4 \) was selected to implement different values of SNR according to (26). Additive zero-mean Gaussian noise \( n(t) \) was also added with \( \sigma_n = 0.1. \) Sampling was performed such that 10 periods of the fundamental component were acquired, i.e., a 200 ms measurement window was taken. The selection of the number of samples \( N_5 \) defines the corresponding sampling frequency. By assuming \( N_5 = 2^{10} \) the corresponding sampling frequency was \( f_s = 5.12 \text{ kHz}, \) and the related frequency resolution was \( \Delta f = 5 \text{ Hz}. \) Three different windows were used, i.e., the rectangular, Hann, and minimum 4-term Blackman-Harris windows. Repeated run analyses (10^4 runs to estimate each average value) were performed to validate the analytical results.

In Fig. 2 the numerical estimates (dotted lines) of the sine-wave detection probability as a function of the false alarm probability are compared with the analytical result (27) (solid lines). The sine wave has SNR = 4 as defined in (26), and a rectangular window (\( SL = 0.637 \)) was used. The normalized frequency fluctuation was \( \Delta x = 0.1, \) and three different values for the mean value of the normalized frequency deviation \( x_0 \) from the closest frequency bin were considered, i.e., 0, 0.25, and 0.5. Clearly by increasing \( x_0 \) the detection probability decreases since the main lobe of the rectangular window in the frequency domain results in the attenuation of the sine wave amplitude according to (12). In Figs. 3 and 4 the same quantities are represented for the Hann (\( SL = 0.849 \)) and the minimum 4-term Blackman-Harris (\( SL = 0.909 \)) windows, respectively. As it was expected, by considering windows with larger SL results in higher detection probability especially...
when the sine wave frequency fluctuates around the edges of the window main lobe, i.e., for $x_0 = 0.5$. Figs. 5 and 6 are similar to Fig. 3 (i.e., Hann window) but with SNR equal to 3 and 2, respectively. According to (26), decreasing SNR results in a lower sine-wave amplitude and larger impact of additive noise. This results in lower detection probabilities (i.e., the curves in Figs. 5 and 6 are lower than curves in Fig. 3), and lower impact of the main lobe of the window in the frequency domain (i.e., the curves in Fig. 5 are close each other more than the curves in Fig. 3, and this phenomenon is further emphasized in Fig. 6 where the behavior is mainly related to the additive noise). In Fig. 7 the Hann window was considered again (and SNR = 4 as in Fig. 3), but the frequency fluctuation $\Delta x$ was increased from 0.1 to 0.5. As it was expected, by increasing the frequency fluctuation the detection probability decreases for $x_0 = 0$ and increase for $x_0 = 0.5$ due to the main lobe of the window in the frequency domain. Therefore, the spread between the three curves for different frequency mean value $x_0$ decreases.

Figs. 8 and 9 compare numerical (dotted lines) and analytical (solid lines) behavior of the mean value and standard deviation of the measured normalized sine-wave amplitude $z$ according to (31) and (32), respectively. Mean value and standard deviation of $z$ are reported as functions of the mean value $x_0$ of the normalized frequency fluctuation. The interval of frequency fluctuation was assumed $\Delta x = 0.5$, and the sine wave amplitude was selected such that SNR = 4. Three pairs of curves are represented corresponding to the three windows already used above. From Fig. 8, as it was expected, the mean value of $z$ is larger for a window with larger SL. Again, this can be explained by the weight provided by the window main lobe. On the contrary, in Fig. 9 it is clear that larger standard deviation of the measured sine-wave amplitude corresponds to windows with lower SL because in this case the window main lobe provides larger variability in weighting the sine wave amplitude. From Figs. 8 and 9 it is clear that even in the case of a properly detected sine wave, both the mean value and the standard deviation can be significantly affected by the sine-wave frequency fluctuation combined with the effect of the window used against spectral leakage.
Fig. 4. Detection probability as a function of the false alarm probability for a sine wave component with normalized frequency fluctuation 0.1 and three different values (i.e., 0, 0.25, and 0.5) for the mean value of the normalized frequency deviation from the closest frequency bin. The SNR (according to (26)) is equal to 4, and the minimum 4-term Blackman-Harris window ($SL = 0.909$) was used.

Fig. 5. Same as Fig. 3 but with SNR = 3.

Fig. 6. Same as Fig. 3 but with SNR = 2.

Fig. 7. Detection probability as a function of the false alarm probability for a sine wave component with normalized frequency fluctuation 0.5 and three different values (i.e., 0, 0.25, and 0.5) for the mean value of the normalized frequency deviation from the closest frequency bin. The SNR (according to (26)) is equal to 4, and the Hann window ($SL = 0.849$) was used.

Fig. 8. Mean value of the normalized sine-wave amplitude $z$ as a function of the normalized mean value $x_0$ of the frequency fluctuation. The normalized frequency interval of fluctuation is $\Delta x = 0.5$, and the sine wave amplitude is such that SNR = 4. Comparison between numerical and analytical results are shown for the three different windows already considered above.

Fig. 9. Standard deviation of the normalized sine-wave amplitude $z$ as a function of the normalized mean value $x_0$ of the frequency fluctuation. The normalized frequency interval of fluctuation is $\Delta x = 0.5$, and the sine wave amplitude is such that SNR = 4. Comparison between numerical and analytical results are shown for the three different windows already considered above.

VII. REFERENCES

