

# STATISTICAL PROPERTIES OF A MODIFIED WELCH METHOD THAT USES SAMPLE PERCENTILES

*Felix Schwock, Shima Abadi*

Department of Electrical and Computer Engineering  
University of Washington, Seattle, WA 98195, USA

## ABSTRACT

We present and analyze an alternative, more robust approach to the Welch's overlapped segment averaging (WOSA) spectral estimator. Our method computes sample percentiles instead of averaging over multiple periodograms to estimate power spectral densities (PSDs). Bias and variance of the proposed estimator are derived for varying sample sizes and arbitrary percentiles. We have found excellent agreement between our expressions and data sampled from a white Gaussian noise process.

**Index Terms**— Spectral estimation, Estimation variance, Welch method

## 1. INTRODUCTION

The Welch's overlapped segment averaging (WOSA) method, first introduced by Welch in 1967 [1], is a popular approach for estimating power spectral densities (PSDs) of stochastic signals due to its computational efficiency, its ability to scale estimation variance, and its potential to reduce spectral leakage. However, the method can suffer from strong outliers in the data caused by transients or other broadband interfering signals. Those outliers can prohibit an accurate estimation of the prevailing noise level, thus, limiting the scope of the WOSA estimator [2]. A possible solution, which has proven to be successful in several spectral estimation applications (see for example [3, 4, 5, 6]), is to take the median of the periodograms at each frequency bin instead of the arithmetic mean. This can be regarded as a special case of a more general Welch estimator that uses sample percentiles, in the following referred to as Welch percentile (WP) estimator. While the statistical properties of the classical WOSA estimator have been analyzed thoroughly [1, 7, 8], respective results for the percentile estimator are yet to be determined. In this paper, we fill this gap by deriving formulas for bias, variance, and limiting distribution of the WP estimator. Section 2 briefly reviews the concept of the classical WOSA estimator and defines the

WP estimator. In Section 3, we derive the statistical properties of the WP estimator. Section 4 compares the theoretical results with Monte Carlo simulations using data sampled from a white Gaussian noise process.

## 2. WELCH PERCENTILE ESTIMATOR

To compute PSD estimates using Welch's method, the time domain signal sampled from a stationary process with sampling frequency  $f_s$  is first divided into  $K$  potentially overlapping segments, each of which containing  $N_s$  samples. Each segment is then multiplied with a window function and the magnitude squared of their fast Fourier transform is computed. The result is a set of modified periodograms  $\{\hat{P}_i(f_j)\}_{i=1}^K$ . Therein,  $f_j$  refers to the  $j$ 's Fourier frequency given by  $f_s/N_s$  and  $\hat{\cdot}$  indicates that each  $\hat{P}_i(f_j)$  is an estimate of some true PSD  $P(f_j)$ . Finally, to obtain the standard WOSA estimate, the average of the modified periodograms is computed at each frequency  $f_j$ .

In contrast to the WOSA estimator, the WP estimator computes the  $q^{\text{th}}$  sample quantile of the set  $\{\hat{P}_i(f_j)\}$  for each  $f_j$  (which is equivalent to the  $p = q \cdot 100$  percentile). To do so, first the order statistic  $\{\hat{P}_{(1)}, \dots, \hat{P}_{(K)}\}$  is determined at each frequency bin. (We have dropped the dependence on  $f_j$  for the sake of brevity.) Afterwards, the  $q^{\text{th}}$  sample quantile can be computed according to [9] as

$$\hat{Q}(q) = K \left( \frac{i}{K} - q \right) \hat{P}_{(i-1)} + K \left( q - \frac{i-1}{K} \right) \hat{P}_{(i)}$$

for  $\frac{i-1}{K} \leq q \leq \frac{i}{K}$  and  $i = 1, \dots, K$ . (1)

That is, if the desired quantile falls between two samples  $\hat{P}_{(i-1)}$  and  $\hat{P}_{(i)}$ , the sample quantile is estimated via linear interpolation. As we will show in Section 3, the sample quantile is, in general, biased compared to the true PSD, whereby the bias  $b$  depends on  $q$  and  $K$ . Hence, the final WP estimator can be defined as

$$\hat{P}_q^{(\text{WP})} = \frac{\hat{Q}(q)}{b(q, K)}. \quad (2)$$

This research was supported by the Office of Naval Research grant number N00014-19-1-2644.

### 3. STATISTICAL PROPERTIES OF THE WP ESTIMATOR

#### 3.1. Distribution

The statistical properties of the WP estimator can be derived from the order statistics  $\{\hat{P}_{(1)}, \dots, \hat{P}_{(K)}\}$  of the modified periodograms. Here, it is assumed that the  $\hat{P}_i$ 's are independent and identically distributed<sup>1</sup>. It is well known (for example, see [8, p. 224-225]) that for a proper window and large enough  $N_s$  the distribution of  $\hat{P}_i$  is given by

$$\hat{P}_i \stackrel{d}{=} \frac{P}{2} \chi_2^2 \quad \text{for } 0 < f_j < \frac{f_s}{2} \quad (3)$$

where  $\chi_2^2$  is the chi-square distribution with two degrees of freedom and probability density function (PDF)

$$f(u) = \begin{cases} \frac{1}{2} e^{-u/2}, & u \geq 0 \\ 0, & u < 0 \end{cases}. \quad (4)$$

According to [10], the PDF  $f_{(i)}(x)$  of the  $i^{\text{th}}$  order statistic  $\hat{P}_{(i)}$  is given by

$$f_{(i)}(x) = \frac{1}{B(i, K-i+1)} F^{i-1}(x) (1-F(x))^{K-i} f(x), \quad (5)$$

where  $F(x)$  is the cumulative distribution function of  $\hat{P}_i$  and can be obtained by integrating Equation (4) from 0 to  $x$ .  $B(\alpha, \beta)$  is the beta function defined by

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt. \quad (6)$$

Equation (5) can now be used to derive expressions for bias and variance of the WP estimator.

#### 3.2. Bias

As shown in [3] for  $q = 0.5$  (i.e., the sample median) and for odd  $K$  the bias can be expressed as

$$b = \sum_{k=1}^K \frac{(-1)^{k+1}}{k}. \quad (7)$$

Following their procedure, similar expressions for an arbitrary quantile and sample size can be derived. To do so, we first substitute  $\alpha = K - i$ ,  $\beta = i - 1$ , and  $t = 1 - F(x)$  and then compute the expected value of  $\hat{Q}(q)$  using Equation (5). We also make the assumption that  $\hat{Q}(q) \approx \hat{P}_{(i)}$  for some  $i = 1, \dots, K$ . While this, in general, does not reflect the WP estimator defined in Equation (1) and (2), we have found that this approximation provides good results for the estimator's statistical properties. The resulting  $\mathbb{E}\{\hat{Q}(q)\}$  is given in Equation (8).

<sup>1</sup>The independence condition holds if adjacent data segments do not overlap, or a moderate overlap along with a proper data taper is used.

$$\mathbb{E}\{\hat{Q}(q)\} \approx -\frac{P}{B(\alpha+1, \beta+1)} \int_0^1 t^\alpha (1-t)^\beta \ln(t) dt. \quad (8)$$

Here, we have used the fact that  $dt = -f(x)dx$  and  $x = -P \ln(1 - F(x)) = -P \ln(t)$  for the chi-square distribution with 2 degrees of freedom. By noting that

$$\frac{\partial t^\alpha (1-t)^\beta}{\partial \alpha} = t^\alpha (1-t)^\beta \ln(t), \quad (9)$$

Equation 8 can be written as

$$\mathbb{E}\{\hat{Q}(q)\} = -\frac{P}{B(\alpha+1, \beta+1)} \frac{\partial}{\partial \alpha} B(\alpha+1, \beta+1). \quad (10)$$

Using the digamma function  $\psi$  to express the partial derivative of the beta function<sup>2</sup>, Equation (10) can be simplified to

$$\mathbb{E}\{\hat{Q}(q)\} = P [\psi(\alpha + \beta + 2) - \psi(\alpha + 1)]. \quad (11)$$

This shows that the bias between the WP estimator and the true PSD  $P$  is given by

$$b = \psi(\alpha + \beta + 2) - \psi(\alpha + 1). \quad (12)$$

Using the fact that  $\psi$  can be expressed as

$$\psi(n) = -\gamma + \sum_{k=1}^{n-1} \frac{1}{k}, \quad \text{for } n \geq 2 \quad (13)$$

where  $\gamma$  is the Euler-Mascheroni constant [11], the bias takes the form of a truncated harmonic series:

$$b = \sum_{k=\alpha+1}^{\alpha+\beta+1} \frac{1}{k}, \quad \text{for } \alpha, \beta \in \mathbb{N}. \quad (14)$$

To express the bias by means of  $K$  and  $q$ , it is helpful to interpret  $\alpha$  and  $\beta$  as the number of samples with values greater and smaller than the desired percentile  $\hat{P}_{(i)}$ . Therefore, we have to distinguish between two cases: (1) the sample percentile  $\hat{Q}(q)$  matches exactly with one of the periodograms  $\hat{P}_{(i)}$  (e.g., if  $q = 0.5$  and  $K$  is odd), or (2) the sample percentile falls in between two periodograms  $\hat{P}_{(i-1)}$  and  $\hat{P}_{(i)}$  (e.g., if  $q = 0.5$  and  $K$  is even). In the former case,  $\alpha$  and  $\beta$  can be expressed by  $\alpha = (K-1)(1-q)$  and  $\beta = (K-1)q$ , respectively. For the latter case,  $\alpha = K(1-q)$  and  $\beta = Kq$  are natural choices. Using this parameterization, Equation (14) can be rewritten as

$$b = \begin{cases} \sum_{k=(K-1)(1-q)+1}^K \frac{1}{k}, & \hat{Q}(q) = \hat{P}_{(i)} \\ \sum_{k=K(1-q)+1}^{K+1} \frac{1}{k}, & \hat{P}_{(i-1)} < \hat{Q}(q) < \hat{P}_{(i)} \end{cases} \quad (15)$$

<sup>2</sup> $\frac{\partial B(\alpha, \beta)}{\partial \alpha} = B(\alpha, \beta) [\psi(\alpha) - \psi(\alpha + \beta)]$

In the limit, that is, for  $K \rightarrow \infty$  both cases converge to  $-\ln(1-q)$ . Furthermore, the products  $(K-1)(1-q)$  and  $K(1-q)$  have to be integers, or, otherwise, rounded to the next nearest integer to compute the bias. If the constellation of  $K$  and  $q$  does not result in an integer value, the polynomial approximation for the digamma function [11]

$$\psi(n) \approx \ln(n) - \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} - \frac{1}{252n^6} \quad (16)$$

can be used to avoid rounding. In this case, the bias should be computed by

$$b = \psi(K+2) - \psi(K(1-q)+1). \quad (17)$$

### 3.3. Variance

In analogy to Equation (8), the second order moment of the sample quantile is given by

$$\mathbb{E}\{\hat{Q}_{(q)}^2\} = \frac{P^2}{B(\alpha+1, \beta+1)} \int_0^1 t^\alpha (1-t)^\beta [\ln(t)]^2 dt. \quad (18)$$

By using the relation

$$\frac{\partial^2 t^\alpha (1-t)^\beta}{\partial \alpha^2} = t^\alpha (1-t)^\beta [\ln(t)]^2 \quad (19)$$

and taking the second derivative of the beta function with respect to  $\alpha$ , the second order moment yields

$$\mathbb{E}\{\hat{Q}_{(q)}^2\} = P^2 \left( [\psi(\alpha+1) - \psi(\alpha+\beta+2)]^2 + [\psi_1(\alpha+1) - \psi_1(\alpha+\beta+2)] \right). \quad (20)$$

Therein,  $\psi_1(n)$  is the derivative of the digamma function – also referred to as trigamma function – and can be approximated by means of Equation (16) as

$$\psi_1(n) = \frac{d\psi(n)}{dn} \approx \frac{1}{n} + \frac{1}{2n^2} + \frac{1}{6n^3} - \frac{1}{30n^5} + \frac{1}{42n^7}. \quad (21)$$

From the first and second order moments of the sample quantile, the variance of the WP estimator can be determined:

$$\text{var}\{\hat{P}_q^{(\text{WP})}\} = \frac{P^2}{b^2} [\psi_1(\alpha+1) - \psi_1(\alpha+\beta+2)]. \quad (22)$$

For the general case, i.e., if  $\hat{P}_{(i)} \neq \hat{Q}(p)$ , the WP estimator's variance can be computed by

$$\text{var}\{\hat{P}_q^{(\text{WP})}\} = \frac{P^2}{b^2} [\psi_1(K(1-q)+1) - \psi_1(K+2)]. \quad (23)$$

### 3.4. Limiting Distribution

For  $K \rightarrow \infty$ , the order statistic of the modified periodograms is normally distributed around  $-P \ln(1-q)$  with variance

$$\text{var}\{\hat{Q}(q)\} = \left(\frac{P}{2}\right)^2 \cdot \frac{q(1-q)}{K f^2(-2 \ln(1-q))}, \quad (24)$$

where  $f$  is the PDF given in Equation (4) [12]. Simplifying this expression and taking the bias correction into account, the limiting variance of the WP estimator can be computed by

$$\text{var}\{P_q^{(\text{WP})}\} = \left(\frac{P}{b}\right)^2 \cdot \frac{q}{K(1-q)}, \quad (25)$$

### 3.5. Equivalent Degree of Freedom

So far, we have assumed that adjacent periodograms are approximately independent. Now, we want to relax this condition by introducing the concept of equivalent degree of freedom (EDOF) to get the number of independent random variables of the quantile estimation. According to [8, p. 429], the EDOF  $\nu$  for the WOSA estimator is

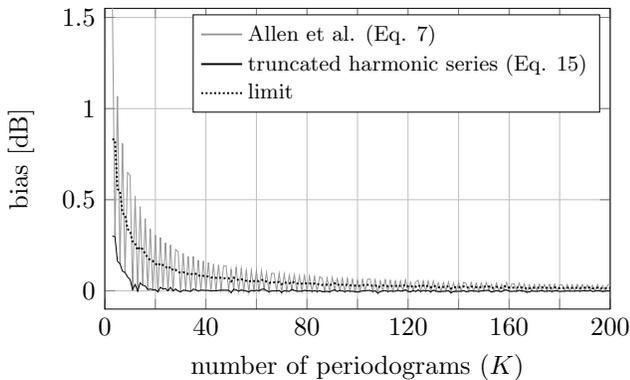
$$\nu = \frac{2K}{1 + 2 \sum_{m=1}^{K-1} \left(1 - \frac{m}{K}\right) \left| \sum_{t=0}^{N_s-1} h_t h_{t+mN_o} \right|}, \quad (26)$$

where  $h_t$  is the data taper and  $N_o$  is the number of overlapping samples. Since the same periodograms are used for the WOSA and WP estimator, Equation (26) also holds for the latter one. That is, the WP estimator uses  $\nu/2$  equivalent and independent periodograms to estimate the true PSD. Bias, variance, and limiting variance can now be computed for arbitrary overlaps and data taper when  $K$  is replaced by  $\nu/2$ . (Note that in Equation (15),  $\nu/2$  and the product  $\frac{\nu}{2}(1-q)$  would need to be rounded to the next nearest integers.)

## 4. SIMULATIONS

Here, we will compare the previously derived expressions for bias and variance with results from a simulated white Gaussian noise sequence. All data segments have a length of  $N_s = 1024$  and a Hann data taper with 50% overlap is used. Subsequently, the WP estimate according to Equation (1) and (2) is computed for various  $q$  and  $K \geq 3$ , and sampling bias and variance are calculated. To reduce the variability in the estimate, 51 100 independent trials are averaged for each  $K$  and  $q$ . Furthermore, the EDOF instead of  $K$  is used in all formulas. It is noted that the goodness of fit between simulations and theoretical results is independent of the data taper if the EDOF is used. This has been tested using the Slepian, Parzen, and triangular window.

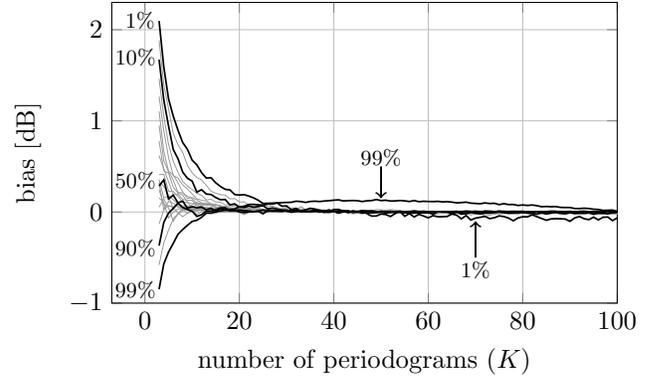
Figure 1 shows the the bias of the Welch 50<sup>th</sup> percentile estimator after applying the bias correction using Equation (7), (15), and the limit  $b = -\ln(0.5)$ . If  $\nu/2$  rounded to the next nearest integer is odd, Equation (7) and (15) give identical results with bias values smaller than 0.1 dB for  $K \geq 7$ . However, if  $\nu/2$  is even, only Equation (15) is capable of accurately compensating the bias. Note that the rounded  $\nu/2$  is in general not equal to  $K$  for the given window and overlap. When using the limit of the bias ( $b = -\ln(0.5)$ ) equally good results for even and odd  $K$  are obtained, but the performance is worse compared to Equation (15). In general the truncated harmonic series is favorable as it gives the lowest bias over all  $K$ . However, for sufficiently large  $K$ , accurate results can be achieved for all three bias correction expressions.



**Fig. 1.** Bias of the Welch 50<sup>th</sup> percentile estimator after correcting the quantile bias according to Equation (7) (Allen et al.), (15) (truncated harmonic series), and  $b = -\ln(0.5)$  (limit).

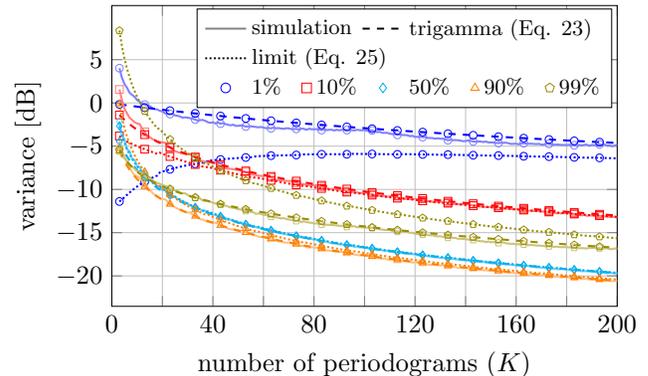
The bias of the WP estimator for different percentiles is shown in Figure 2. Therein, the digamma approximation (Equation (16) and (17)) is used to compensate for the quantile bias. One can observe that, for small values of  $K$ , more extreme percentiles tend to over or underestimate the true PSD, whereas percentiles around 63% (unbiased estimator) exhibit only a small or no bias. Only for the 1<sup>st</sup> and 99<sup>th</sup> percentile a bias greater than 0.1 dB can still be observed for some  $K \geq 30$ . (This bias will also vanish as  $K$  further increases.)

Finally, variance and limiting variance according to Equation (23) and (25) are compared to the sampling variance of the WP estimator in Figure 3. The bias is corrected using the digamma approximation. The results show that Equation (23) deviates by less than 0.5 dB from the simulations for  $K \geq 16$  and percentiles between 10% and 90%. The limiting variance (Equation (25)), on the other hand, requires values  $K \geq 79$  to provide the same accuracy. For the 1<sup>st</sup> and 99<sup>th</sup> percentile, a greater deviation between theoretical expressions and simulations can be observed. In these cases,



**Fig. 2.** Bias of the WP estimator for percentiles between 1% and 99% in 5% increments after correcting the quantile bias using the digamma approximation (Equation (16) and (17)).

larger values of  $K$  would be necessary to achieve a better fit. Figure 3 also shows that the variance of the 50<sup>th</sup> percentile estimator (median) is larger compared to the variance of the 90<sup>th</sup> percentile estimator. Indeed one can show that, in the limit, the 80<sup>th</sup> percentile estimator has the lowest variance – by a factor of approximately 1.3 dB compared to the median – among all WP estimators.



**Fig. 3.** Simulated and theoretical variance according to Equation (23) (trigamma) and Equation (25) (limit) of the WP estimator.

## 5. CONCLUSION

The WP estimator is a robust approach for computing PSD estimates. Equations for the bias of the underlying quantile estimate have been derived. We have shown that our bias correction approach outperforms the existing method for the Welch 50<sup>th</sup> percentile estimator and also performs excellent for other percentiles. Furthermore, simple expressions for the estimator's variance have been derived and comparisons with simulated data have shown great agreement for most percentiles and a wide range of sample sizes.

## 6. REFERENCES

- [1] P. Welch, "The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.
- [2] R. A. Maronna, R. Martin, V. J. Yohai, and M. Salibián-Barrera, *Robust Statistics, 2nd Edition*, Wiley, 2nd edition, 2019.
- [3] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, "Findchirp: An algorithm for detection of gravitational waves from inspiraling compact binaries," *Physical Review D*, vol. 85, no. 12, Jun 2012.
- [4] S. E. Parks, I. Urazghildiiev, and C. W. Clark, "Variability in ambient noise levels and call parameters of north atlantic right whales in three habitat areas," *The Journal of the Acoustical Society of America*, vol. 125, no. 2, pp. 1230–1239, 2009.
- [5] N. D. Merchant, P. Blondel, D. T. Dakin, and J. Dorocicz, "Averaging underwater noise levels for environmental assessment of shipping," *The Journal of the Acoustical Society of America*, vol. 132, no. 4, pp. EL343–EL349, 2012.
- [6] J. T. Kees, J. M. Ernst, W. C. Headley, and A. A. Louis Beex, "Robust blind spectral estimation in the presence of non-gaussian noise," in *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, Baltimore, MD, USA, 2017, pp. 629–634.
- [7] A. H. Nuttall, "Spectral estimation by means of overlapped fast fourier transform processing of windowed data," Tech. Rep. 4169, NUSC, New London, CT, USA, 1971.
- [8] D. B. Percival and A. T. Walden, *Spectral Analysis for Univariate Time Series*, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, UK, 2020.
- [9] E. Parzen, "Nonparametric statistical data modeling," *Journal of the American Statistical Association*, vol. 74, no. 365, pp. 105–121, 1979.
- [10] H. A. David and H. N. Nagaraja, "Distribution of a single order statistic," in *Order Statistics*, pp. 9–11. Wiley, 3rd edition, 2004.
- [11] M. Abramowitz and I. Stegun, "Psi (digamma) function," in *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Applied mathematics series (Washington, D.C.); 55, p. 258. U.S. Dept. of Commerce : U.S. G.P.O., Washington, D.C., USA, 10th print., with corrections. edition, 1972.
- [12] M. G. Kendall, A. Stuart, J. K. Ord, S. F. Arnold, and A. O'Hagan, "Standard errors of quantiles," in *Kendall's advanced theory of statistics*, Kendall's library of statistics, pp. 356–358. Edward Arnold, London, UK, 6th edition, 1994.