Fast and Accurate Power Spectral Analysis of Heart Rate Variability using Fast Gaussian Gridding

Charalampos Eleftheriadis, Georgios Karakonstantis

Institute of Electronics, Communication, and Information Technology (ECIT)
Queen's University Belfast, United Kingdom

Abstract

In this paper, we propose an algorithm for fast and accurate Power Spectral Analysis of Heart Rate Variability based on the Lomb Periodogram. The previously introduced Fast-Lomb periodogram, may have reduced the computational complexity of PSA, however it still requires a large oversampling factor, which increases the complexity of the needed FFTs. In our approach, by utilising the Fast Gaussian Gridding method we produce accurate evenly spaced grids for the required FFTs by restricting the oversampling factor only to 2. By doing so, the required FFT size is reduced by up to 4 times without compromising the output accuracy. Our results indicate that the proposed spectral analysis system can achieve up-to 76.55% savings in the number of operations or up-to 75.8% in terms of the total execution time.

1. Introduction

In the last decade, the widespread usage of Wireless Body Sensor Networks (WBSNs) and bio-signal monitoring allowed the realization of smart health applications that can help in preventing fatal conditions and diseases [1]. One of the most important indicators of a subject's health is Heart Rate Variability (HRV), which is exploited by many emerging health monitoring applications, since heart conditions are responsible for nearly one third of all global deaths [2].

HRV is extracted via the processing of specific Electrocardiogram (ECG) features and various methods in time or frequency domain were introduced in the past years to estimate it. Power Spectral Analysis (PSA) of HRV can provide more insights about underlying conditions, (compared with the conventional time domain methods), especially when applied on short term time windows [3], however, its efficient computation is extremely challenging. In particular, during the delineation of the ECG, the reference points (R peaks) are unevenly distributed since interbeat intervals vary; thereby, the traditional low complexity Fast Fourier Transform (FFT) cannot be used directly to compute the Power Spectral Density (PSD). Usually, to tackle this problem an interpolation method is used to construct a regular grid and after that the PSD is estimated via a FFT transform or an autoregressive (AR) model. However, such methods may alter the spectral contents of the signal as the applied interpolation has a low pass filtering effect [4].

For this reason, the Lomb-Scargle periodogram has been introduced for the PSA of HRV, which is based on the least-squares fitting of the selected RR intervals at the specified frequencies with complexity $O(N^2)$. Fortunately, the Fast-Lomb algorithm [4], reduced the complexity by approximating the required trigonometric sums with two Lagrange based interpolated time-series followed by two FFTs as shown in Fig 1.

However, even in this case a high oversampling factor is used during the interpolation stage, which leads to two computationally expensive FFTs that are prohibiting the realization of such an analysis on portable devices. Recent approaches tried to address this by approximating the FFTs with Wavelets [5] however, such a method come at cost of accuracy loss and does not address the problem of the default, large oversampling factor.

In this paper we propose to produce a more accurate grid by employing the Fast Gaussian Gridding (FGG) algorithm [6] during the interpolation phase. By doing so we can maintain high accuracy with a smaller oversampling factor, resulting in a reduced FFT size. The contributions of the paper can be summarized as follows:

1) Analyze the main stages of the Fast-Lomb method and reveal the disadvantages of the Lagrange based resampling method in terms of performance.
2) Develop a Fast-Lomb algorithm based on the Fast Gaussian Gridding method (FGG), and produce an accurate evenly spaced grid.
3) Show the achievable savings of the proposed method in terms of operations and time that can reach up to 76.55% and 75.8%, respectively. Such savings come at no accuracy loss in PSD output when compared with the traditional Fast Lomb algorithm for numerous cardiac samples.

The rest of the paper is organized as follows. Section 2 describes the background of the conventional Lomb and

Fast-Lomb algorithms. Section 3 presents our proposed approach; the Fast Gaussian Gridding algorithm applied on the Fast-Lomb method, 4 describes the achievable savings by FFT size reduction. Finally, conclusions are drawn in Section 5.

2. Background

In this section we initially give a brief overview of the existing algorithms for the calculation of the Lomb periodogram, and we highlight the main principles of the Fast Gaussian gridding.

2.1 PSA on unevenly spaced data

The extracted RR-tachogram from the ECG is comprised of an unevenly spaced time series so for the evaluation of the PSD the Lomb-Scargle periodogram can be employed. Generally, the power $PSD(\omega)$ is:

$$PSD(\omega) = \frac{1}{2}\left(\frac{\left[\sum_j X_j \cos\omega(t_j-\tau)\right]^2}{\sum_j \cos^2\omega(t_j-\tau)} + \frac{\left[\sum_j X_j \sin\omega(t_j-\tau)\right]^2}{\sum_j \sin^2\omega(t_j-\tau)}\right) \quad (1)$$

ω: angular frequency, t_j: observation time,

X_j: Detrend RR interval,

τ: constant offset for invariant to time shift PSD.

So, using some trigonometric identities, for each frequency ω we have to compute four sums as shown in Eq. 2, 3:

$$S_h(\omega) = \sum_j X_j \sin(\omega t_j),\ \ C_h(\omega) = \sum_j X_j \cos(\omega t_j) \quad (2)$$

$$S_2(\omega) = \sum_j \sin(2\omega t_j),\ \ C_2(\omega) = \sum_j \cos(2\omega t_j) \quad (3)$$

Then we compute the four sums of Eq. 1 as:

$$\sum_j X_j \cos\omega(t_j-\tau) = C_h(\omega)\cos(\omega\tau) + S_h(\omega)\sin(\omega\tau) \quad (4)$$

$$\sum_j X_j \sin\omega(t_j-\tau) = S_h(\omega)\cos(\omega\tau) - C_h(\omega)\sin(\omega\tau) \quad (5)$$

$$\sum_j \cos^2\omega(t_j-\tau) = \frac{N}{2} + \frac{1}{2}C_2(\omega)\cos(2\omega\tau) + \frac{1}{2}S_2(\omega)\sin(2\omega\tau) \quad (6)$$

$$\sum_j \sin^2\omega(t_j-\tau) = \frac{N}{2} - \frac{1}{2}C_2(\omega)\cos(2\omega\tau) - \frac{1}{2}S_2(\omega)\sin(2\omega\tau) \quad (7)$$

The computation of the PSD for M (usually $N = M$) number for output frequencies using the Eq.1 can be very expensive as $O(N)$ operations are required to compute $PSD(\omega_i)$, thus the total complexity is $O(N^2)$.

The Fast-Lomb method reduces the complexity by computing the sums of Eq. 2, 3 using a Lagrange based resampling followed by two FFTs. So, from Eq. 2:

$$X_h(\omega) = C_h(\omega) - i * S_h(\omega) = \sum_j X_j e^{-i\omega t_j} \quad (8)$$

We can write the exponential term as:

$$g(t) = e^{-i\omega t} \approx \sum_k w_k(t) e^{-i\omega \widehat{t_k}} \quad (9)$$

Where $w_k(t)$ is the kth order Langrage polynomial and $\widehat{t_k}$ are evenly spaced time-series points. If we place Eq. 9 in Eq. 8, we have:

$$X_h(\omega) = \sum_j X_j g(t_j) \approx \sum_j X_j \sum_k w_k(t_j) e^{-i\omega\widehat{t_k}} =$$

$$= \sum_k \sum_j X_j w_k(t_j) e^{-i\omega\widehat{t_k}} = \sum_k \hat{h}_k e^{-i\omega\widehat{t_k}} \quad (10)$$

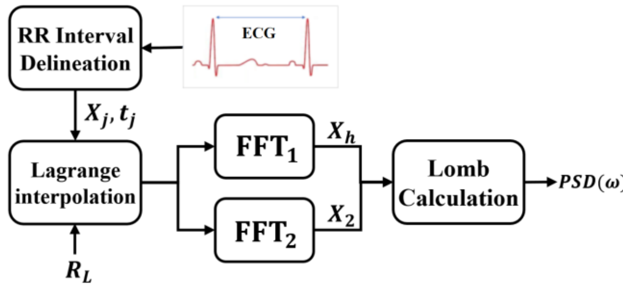

Fig. 1: The block diagram of the conventional Fast-Lomb periodogram; The ECG delineation is followed by a Lagrange based resampling algorithm succeeded by the X_h and X_2 computation via 2 FFTs. These FFTs do not compute the PSD, as their only purpose is to lower the complexity the Lomb periodogram.

So, by knowing $\hat{h}_k = \sum_j X_j w_k(t_j)$ we can use and FFT to computed the $X_h(\omega)$. In order adequately approximate Eq. 9, we have to oversample the number of the output frequencies by a factor of R_L. Furthermore, employing an FFT in Eq. 10 will result in the double-sided spectrum of X_h. We can follow the same procedure by replacing $X_j = 1$ and by doubling the frequency ω to 2ω, to compute $X_2(\omega) = C_2(\omega) - i * S_2(\omega)$. Some typical oversampling factors are $R_L = 4, 8$ thus according to [4] for N output frequencies the size of the FFT will be $4R_LN$. The block diagram of the Fast-Lomb algorithm is depicted in Fig.1. where there are two FFTs blocks ($\text{FFT}_1, \text{FFT}_2$) that are the most computationally expensive parts of the algorithm and they compute X_h, X_2.

2.2 Challenges

In spite of the invention of the Fast-Lomb periodogram, the algorithm has some clear drawbacks.

- For smaller values of N it has higher complexity than the Lomb algorithm as is shown later in our results in Fig. 3.
- The uniform grid produced by the Lagrange interpolation uses a high oversampling factor resulting in an oversized FFT, which leads to large computational overhead.

3 Proposed approach

To address the challenges mentioned in the previous section we introduce an approach that replaces the Lagrange based interpolation with the Fast Gaussian gridding method. By doing this we show that we can construct a more accurate grid, thus restricting the oversampling factor.

3.1 Fast Gaussian Gridding

Initially, before going into the details let us explain the principles of the FGG algorithm. Fast Gaussian Gridding [6] is a method, which we exploit for medical imaging, astronomy and numerical analysis applications similar to Eq. 8. The goal is to compute an output vector $F \in \mathbb{R}^M$ from an input vector $f \in \mathbb{R}^N$. The value of f_j is known at non-equally spaced $x_j \in [0,2\pi)$ points

So:

$$F(k) = \frac{1}{N}\sum_j f_j e^{-ikx_j}, x_j \in [0,2\pi), k \in [-\frac{M}{2}, \frac{M}{2} - 1] \quad (11)$$

From Eq. 11 f_j can be described by a 2π periodic function:

$$f(x) = \sum_j f_j \delta(x - x_j) \quad (12)$$

where $\delta(x)$ is the Dirac delta function. Also, by using a 2π periodic gaussian kernel $g_\tau(x)$:

$$g_\tau(x) = \sum_{l=-\infty}^{\infty} e^{-(x-2l\pi)^2/4\tau} \quad (13)$$

We define the convolution of $f_\tau(x)$ as

$$f_\tau(x) = f(x) * g_\tau(x) = \sum_{j=0}^{N-1} f_j g_\tau(x - x_j)$$

$$= \sum_{j=0}^{N-1} f_j \sum_{l=-\infty}^{\infty} e^{-(x-x_j-2l\pi)^2/4\tau} \quad (14)$$

The kernel parameter τ depends on the approximation points that will be used for the gaussian kernel (Eq.18). Following that, we set $M_r = 2R_F M$ where R_F is an oversampling factor (typically $R_F = 2$). The M_r point FFT is:

$$F_\tau(k) = \frac{1}{M_r}\sum_{m=0}^{M_r-1} f_\tau(2\pi m/M_r) e^{-ik2\pi m/M_r} \quad (15)$$

Therefore, from the convolution theorem:

$$f_\tau(x) = f(x) * g_\tau(x) \Rightarrow F_\tau(k) = F(k)\, G_\tau(k)$$

$$F_\tau(k) = F(k)\, G_\tau(k) \quad (16)$$

Thus, the deconvolution operation is,

$$F(k) = \sqrt{\frac{\pi}{\tau}} e^{k^2\tau} F_\tau(k) \quad (17)$$

Since, $f(x)$ which is known at non evenly spaced points x_j is convoluted with a gaussian kernel $g_\tau(x)$. This produces a uniformly spaced grid at points $2\pi m/M_r, m \in [0, M_r - 1]$. Moreover, the Fast Gaussian Gridding approximation can be used, to lower the complexity of Eq 14.

Table 1: Complexities of the FGG and Langrage methods (where $M = N$) the FFT1 computes the sums of eq 2 and FFT2 computes the sums of eq 3.

Lomb Method	Complexities		
	Grid	FFT1	FFT2
FGG	$O(2M_{sp}N)$	$O(2R_F N log(2R_F N))$	$O(4R_F N log(4R_F N))$
Lagrange	$O(R_L N)$	$O(4R_L N log(4R_L N))$	$O(4R_L N log(4R_L N))$

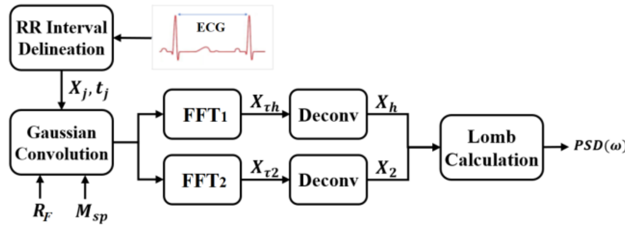

Fig. 2: The block diagram of the FGG based Fast-Lomb periodogram; The ECG delineation is followed by a Gaussian "smearing" algorithm succeeded by the two FFTs and two deconvolutions (Eq. 17).

Therefore, if we only take into account $2M_{sp}$ gaussian points when performing the convolution, thus the Eq 14. is approximated as:

$$f_\tau(x) \approx \sum_{j=0}^{N-1} f_j \sum_{l=-Msp}^{Msp-1} e^{-(x-x_j-2l\pi)^2/4\tau} \quad (18)$$

3.2 Fast-Lomb periodogram via Fast Gaussian Gridding

In our approach, we rescale Eq. 8 to derive Eq. 11. so, we have Eq. 19:

$$X_h(\omega) = X_h(k) = \sum_j X_j e^{-i2\pi k\Delta f t_j}, \quad X_j = f_j \quad (19)$$

$$2\pi\Delta f t_j = x_j \mid t_{N-1} < 1/\Delta f$$

Also, by applying these transforms to $X_2(\omega)$, we can replace the traditional Lagrange based interpolation with the Fast Gaussian Gridding. The block diagram of the Fast-Lomb periodogram via FGG is the one in Fig. 2. During the computation of FFT2, both methods required a $4RN$ size FFT (where $R = R_F$ or R_L) as we have to compute X_2 (eq 3) at 2ω frequencies. Furthermore, regarding the FFT1, in [4] the authors compute a $4RN$ FFT size, whereas in our proposed method we can use a $2RN$ FFT size.

To sum up, the FGG method not only requires a smaller oversampling factor R_F, but also, it allows the reduction of FFT1 by 2. In addition, the factor R_F is smaller than the R_L for the same accuracy of operations. Depending on the sizes N and M of the input and output vectors (in our case $M = N$), they add different overhead to the computation of the PSD which can be summarized in table 1.

4. Results

To determine the computational gains in terms of the number of operations and the total execution time of our proposed method, we used synthetic and real RR interval samples from the MIT-Arrythmia database [7]. For our tests we chose different time windows of fixed input points N and and we set the output frequency points $M = N$. We also compared both methods with the direct computation of the Lomb periodogram and we measured the two-norm error of the outputs.

Specifically, Fig 3a. depicts the increase on total

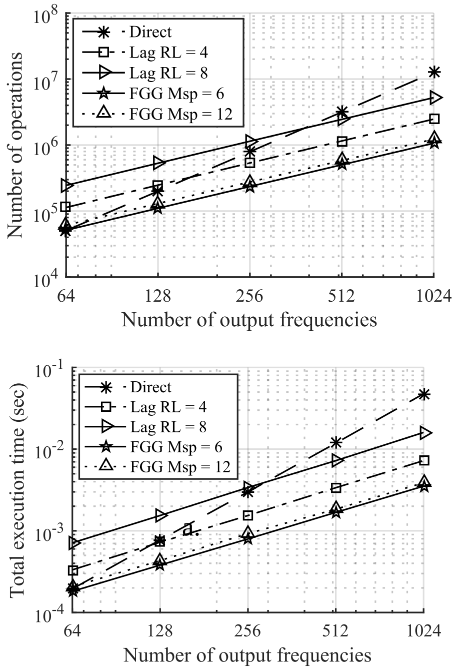

Fig. 3: At a. we compare the number of operations for the *PSD* estimation for different *N* using the; direct method, the conventional Fast-Lomb method and the FGG Fast-Lomb method. At b. we compare them in terms of total execution time.

operations for different numbers of N, while Fig. 3b. compares the methods in terms of total execution time. Regarding the Lagrange based method, doubling the factor R_L(4 to 8) to improve the accuracy, increases on average by 53% the total operations and by 54% the execution time. In contrast, the FGG method presents a slight rise of 15% and 11% in respect of both metrics, as doubling M_{sp} affects only the linear gridding term $O(2M_{sp}N)$. In all cases the oversampling factor R_F is to set 2 leading to smaller FFTs. In addition, due to the high oversampling factor R_L, at the lowest values of N the direct method outperforms the Lagrange based while the FGG is always faster.

To examine the timing gains compared to the two-norm error of each method, we performed 3 gridding tests which achieve similar accuracy for various sizes N (Fig. 4). Regarding the Lagrange based method we chose $R_L = 2, 4, 8$, while for the FGG method we opted for $M_{sp} = 3, 6, 12$. In table 2 we compare the methods in terms of accuracy, timing speedup and we calculate the added overhead for more precise computation of the PSD.

Table 2: Comparison of the achieved accuracy of the two methods.

R_L/M_{sp}	2-norm error	Timing	
		2x Overhead %	Speedup %
2	1.13e-01	-	10.57
3	5.47e-02		
4	1.05e-03	53.48	49.32
6	1.07e-04	18.03	
8	3.64e-09	57.62	74.34
12	2.62e-10	15.96	

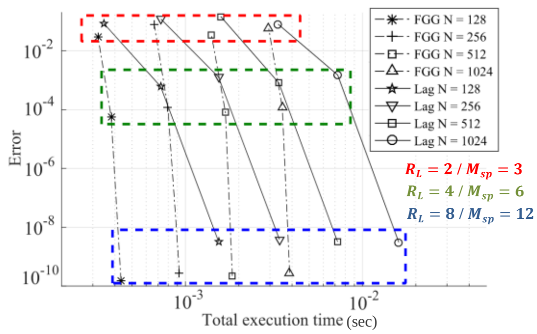

Fig. 4: Conventional Fast-Lomb algorithm versus the FGG Fast-Lomb; There are 3 sets of computations $R_L/M_{sp} = 2/4$, $R_L/M_{sp} = 4/6$ $R_L/M_{sp} = 8/12$.

So, as shown in Fig 4. the FGG method always outperforms the Langrage especially when higher accuracy is required. Not only it achieves a lower 2-norm error but also it is significantly faster, as the gaussian gridding produces a more accurate grid for a fixed size FFT. We observe that the highest speedup is achieved for $M_{sp} = 12$ and $R_L = 8$ for $N = 1024$ where the total number of operations is reduced by 76.55% while the execution time is 75.8% less.

5. Conclusion

In this paper, an alternative algorithm for PSA via the Lomb-Scargle algorithm was presented. It was shown that the proposed approach by utilizing the Fast Gaussian Gridding Method can produce more accurate grids than the conventional method, requiring smaller FFT sizes. Eventually this approach can reduce the required power allowing the implementation of real time PSA of bio-signals on portable devices.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 programme under the Marie Skłodowska-Curie grant agreement No. 945231.

References

[1] P. L. Malasinghe and e. al., "Remote Patient Monitoring: a Comprehensive Study," *JAIHC*, pp. 59-70, 2019
[2] WHO, Cardiovascular diseases (CVDs) [Online].
[3] F. Shaffer, et. al., "An Overview of Heart Rate Variablity Metrics and Norms," *Frontiers in Public Health*, pp. 1-11, 2017.
[4] W. H. Press, "Fast Algorithm for Spectral Analysis of Unevenly Sampled Data," *Astrophysical Journal*, 1989.
[5] G. Karakonstantis,et. al., "A Quality-Scalable and Energy-Efficient Approach for Spectral Analysis of Heart Rate Variability," *DATE*, 2014.
[6] L.Greengard et. al., "Accelerating the NonuniformFast Fourier Transform," *SIAM REVIEW*, vol. 46, p. 443–454, 2003.
[7] MIT-BIH arrhythmia database. [Online].

Contact: C. Eleftheriadis (C.Eleftheriadis@qub.ac.uk) 18 Malone Rd, Belfast, UK