

Statistical Assessment of Performance of Algorithms for Detrending RR Series

Antonio Fasano¹ and Valeria Villani²

Abstract—Detrending RR series is a common processing step prior to HRV analysis. Customarily, RR series, which are inherently unevenly sampled, are interpolated and uniformly resampled, thus introducing errors in subsequent HRV analysis. We have recently proposed a novel approach to detrending unevenly sampled series, which is based on the notion of *weighted quadratic variation reduction*. In this paper, we extensively assess its performance on RR series through a statistical analysis. Numerical results confirm the effectiveness of the approach, which outperforms state-of-the-art methods. Furthermore, it is *statistically uniformly better* than competing algorithms. A sensitivity analysis shows that it is robust to variations of its controlling parameter. The algorithm is simple and favorable in terms of computational complexity, thus being suitable for long-term HRV analysis. To the best of the authors' knowledge, it is the *fastest* algorithm for detrending RR series.

I. INTRODUCTION

Heart rate variability (HRV), that is the variation over time of the period between consecutive heartbeats, is a noninvasive quantitative marker of autonomic activity. It is used to investigate the balance between the sympathetic and parasympathetic branches of the autonomic nervous system [1]. HRV is assessed from RR interval time series, or tachograms, by computing some time or frequency-domain metrics [1]–[3]. HRV analysis is usually preceded by a detrending step of the RR series. This is mainly due to the fact that most frequency-domain metrics require estimation of the power spectral density and slowly-varying trends introduce nonstationarity, which adversely affects subsequent analysis [1]. Further, the analysis of slow trends in short-term HRV is considered a dubious measure [1].

Classical approaches to RR series detrending consist of subtracting the trend estimated using first or higher order polynomials [4], [5]. More recently, a technique based on smoothness priors was proposed in [6]. The major limitation of these approaches lies in the fact that they apply to uniformly sampled series, despite RR series are inherently unevenly sampled. Thus, linear or cubic spline interpolation and resampling are required prior to detrending [2], [5]. However, it has been shown that the interpolation-resampling process introduces significant errors in the spectral analysis of HRV [7], [8]. In particular, it overestimates the total power in the LF and HF bands, and the LF/HF ratio, the most used frequency-domain measures of HRV [2], [7]. Recently, an algorithm, called smoothing by Gaussian process priors, for

detrending RR unevenly sampled series has been proposed in [9].

In our recent paper [10] we have proposed a novel approach to detrend unevenly sampled series, which is based on the notion of *weighted quadratic variation reduction* (WQVR). It does not rely on the interpolation-resampling process. Hence, detrending by WQVR can be followed by spectral analysis using the Lomb-Scargle periodogram [2], which is tailored to unevenly sampled data. Moreover, the algorithm is simple and remarkably fast, thus lending itself to the detrending of 24-hour RR series for long-term HRV analysis [1].

In this paper, we carry out a comparative analysis with state-of-the-art algorithms and provide a complete statistical assessment of detrending performance of our and competing approaches. Numerical results highlight the superiority of detrending by WQVR, which outperforms state-of-the-art algorithms. Further, we introduce the concept of statistically uniformly better algorithm, which characterizes the algorithms that have uniformly better performance over the full range of reconstruction errors, and show that our approach is *statistically uniformly better* than competing algorithms. Finally, a sensitivity analysis shows that it is robust to variations of its controlling parameter.

II. MATERIALS AND METHODS

A. RR series

For a quantitative assessment of the performance of our detrending algorithm and competing approaches, we considered synthetic RR series (trend-free) affected by simulated trends. The synthetic RR series, denoted by $\mathcal{R}^{(0)}$, was generated according to the model described in [11], with LF/HF ratio equal to 0.5, heart rate with mean 60 bpm and standard deviation 5 bpm, and duration 4.5 minutes. Then, $\mathcal{R}^{(0)}$ was corrupted by synthetic trends ξ_i rendered as white Gaussian noise ($\mu = 0$, $\sigma^2 = 1$) low-pass filtered with bandwidth 0.05 Hz, unevenly sampled at the time of the series $\mathcal{R}^{(0)}$. The resulting series is denoted by $\tilde{\mathcal{R}}_i = \mathcal{R}^{(0)} + \xi_i$. We generated 300 independent realizations of synthetic trends ξ_i .

B. Detrending RR series by WQVR

We shortly recall the rationale behind the approach based on WQVR, further details are given in [10]. Trends are low¹ “variability” components of RR series. Thus, provided that we introduce a suitable measure of “variability”, a trend

¹Antonio Fasano is with the Faculty of Engineering, Università Campus Bio-Medico di Roma, Roma, Italy a.fasano@unicampus.it

²Valeria Villani is with the Dipartimento di Scienze e Metodi dell'Ingegneria, Università degli Studi di Modena e Reggio Emilia, Reggio Emilia, Italy valeria.villani@unimore.it

¹Low with respect to RR series “variability”.

can be estimated searching for the low-variability component closest, in some sense, to the RR series. The estimated trend is then subtracted from the RR series.

We quantify the variability of a generic vector introducing the weighted quadratic variation (WQV) [10]. Given a vector $\mathbf{x} = [x_1 \cdots x_n]^T \in \mathbb{R}^n$, with $n \geq 2$, and the set $\mathcal{W} = \{w_1, \dots, w_{n-1}\}$, with $w_k > 0$, the *weighted quadratic variation* of \mathbf{x} , denoted by $[\mathbf{x}]_{\mathcal{W}}$, is defined as

$$[\mathbf{x}]_{\mathcal{W}} \doteq \sum_{k=1}^{n-1} w_k^2 (x_k - x_{k+1})^2 = \|\mathbf{D}_{\mathcal{W}} \mathbf{x}\|^2, \quad (1)$$

where $\|\cdot\|$ is the Euclidean norm and $\mathbf{D}_{\mathcal{W}}$ is the $(n-1) \times n$ matrix with entries

$$[\mathbf{D}_{\mathcal{W}}]_{i,j} = w_i (\delta_{i,j} - \delta_{i+1,j}) \quad (2)$$

where $\delta_{i,j}$ is the Kronecker delta. It is possible to prove that the definition (1) is well posed and the WQV is a consistent measure of variability.

The WQV is well suited for measuring the variability of unevenly sampled series. Indeed, by suitably choosing the set of weights \mathcal{W} it is possible to introduce into the WQV information about the – possibly uneven – sampling grid. Indeed, let $\mathbf{x} = [x_1, \dots, x_n]^T$ be a vector collecting samples of a series taken at (non-uniform) times t_1, \dots, t_n , with $t_i < t_{i+1}$. By setting weights to

$$w_k = \frac{1}{t_{k+1} - t_k}, \quad k = 1, \dots, n-1 \quad (3)$$

$[\mathbf{x}]_{\mathcal{W}}$ becomes a measure of variability that takes into account both values and sampling times.

Now, let us apply these results to RR series

$$RR_k = R_{k+1} - R_k, \quad k = 1, \dots, n \quad (4)$$

where R_k is the time lag of the k th R-peak. The corresponding time instants are

$$t_k = \sum_{i=1}^{k-1} RR_i, \quad k = 1, \dots, n \quad (5)$$

with $t_1 = 0$. Thus, taking into account both (3) and (5), the weights for RR series become

$$w_k = RR_k^{-1}, \quad k = 1, \dots, n-1. \quad (6)$$

In the following, we denote by $\tilde{\mathcal{R}}$ the vector collecting n samples of an RR series, which is affected by a slowly varying trend, by \mathbf{x} the vector of estimated trend, and by \mathcal{R} the corresponding RR series detrended. We propose to estimate the trend \mathbf{x} by solving the following convex optimization problem

$$\begin{cases} \text{minimize} & \|\mathbf{x} - \tilde{\mathcal{R}}\|^2 \\ \text{subject to} & [\mathbf{x}]_{\mathcal{W}} \leq \rho \end{cases} \quad (7)$$

where weights are set according to (6) and ρ is a nonnegative constant that controls the WQV of the estimated trend. It is possible to prove [10] that the solution to (7) is given by

$$\mathbf{x} = \left(\mathbf{I} + \lambda \mathbf{D}_{\mathcal{W}}^T \mathbf{D}_{\mathcal{W}} \right)^{-1} \tilde{\mathcal{R}} \quad (8)$$

where \mathbf{I} denotes the $n \times n$ identity matrix, and $\lambda = \lambda(\rho)$ is a nonnegative parameter that controls the degree of variability of the estimated baseline. It is in one-to-one correspondence with ρ . Once estimated, the trend can be removed from the RR series by subtraction

$$\mathcal{R} = \tilde{\mathcal{R}} - \mathbf{x} = \left[\mathbf{I} - (\mathbf{I} + \lambda \mathbf{D}_{\mathcal{W}}^T \mathbf{D}_{\mathcal{W}})^{-1} \right] \tilde{\mathcal{R}}. \quad (9)$$

It is possible to prove that trend estimation using (8) is *very fast* [10], as its complexity is $O(n)$, i.e., *linear* in the size of the RR series to detrend.

C. Competing algorithms

We compared our algorithm with: i) the smoothness priors (SP) approach [6], which is one of the most used, if not the most used, techniques for RR series detrending; ii) the smoothing by Gaussian process priors (SGP) method [9], which can handle unevenly sampled series. SP was implemented using the code provided in [6]. It can handle only uniformly sampled data. As a consequence, the corresponding detrended series $\mathcal{R}_i^{(\text{SP})}$ was unevenly resampled at the same time of $\mathcal{R}^{(0)}$, to allow a meaningful comparison of performance. For SGP we used the MATLAB script made available by the authors on the website reported in [9].

D. Performance metrics

The quality of detrending was assessed through the following errors

$$\varepsilon_i(\text{alg}_k) = \left\| \mathcal{R}_i^{(\text{alg}_k)} - \mathcal{R}^{(0)} \right\|^2, \quad i = 1, \dots, 300 \quad (10)$$

where $\mathcal{R}_i^{(\text{alg}_k)}$ is the series detrended using algorithm $\text{alg}_k \in \{\text{WQVR}, \text{SP}, \text{SGP}\}$. Performance of different algorithms is measured in terms of the empirical distribution function of the corresponding errors (10), namely

$$\hat{F}_{\text{alg}_k}(\epsilon) = \frac{1}{N} \sum_{i=1}^N \chi_{(-\infty, \epsilon]} [\varepsilon_i(\text{alg}_k)] \quad (11)$$

where $\chi_E[\cdot]$ denotes the indicator function of the set E , and N is the number of generated trend realizations. Note that the empirical distribution function (11) provides a *complete* statistical description of the performance of each algorithm, as opposed to the mean, the variance, or the median, which give only partial information: different algorithms may exhibit errors with the same mean, median, and variance, even though the corresponding empirical distribution functions are completely different.

The use of the empirical distribution function as a performance metric is motivated by the following reasons. Firstly, the Glivenko-Cantelli Theorem [12] guarantees that almost surely as $N \rightarrow \infty$, the function in (11) converges uniformly to the true cumulative distribution function, since the maximum gap between the two of them goes to zero. Moreover, the estimation error in approximating the true cumulative distribution function can be controlled by the Dvoretzky–Kiefer–Wolfowitz inequality [13]. Note that no similar result holds for the histogram, which moreover suffers from some problems that make it not the best choice

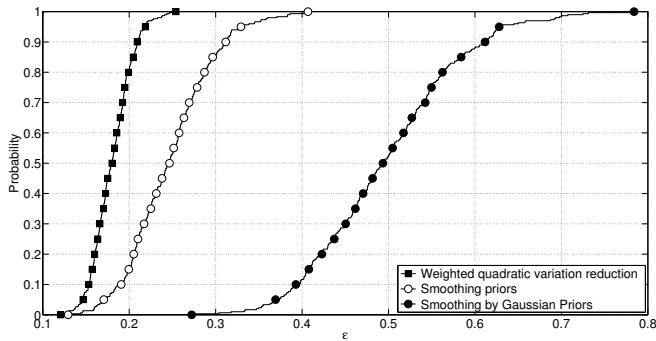


Fig. 1. Empirical distribution functions of the error (10) for WQVR, SP and SGP.

for density estimation [13]. Secondly, an important aspect of empirical distribution functions is that they allow us to compare algorithms over the *full range* of errors taking account of error relative frequencies. This cannot be achieved by considering sample mean and variance only. Moreover, empirical distribution functions are useful for characterizing algorithms that have uniformly better performance over the whole range of errors, where better is meant as statistically better.

To this end, denoting by $\hat{F}_A(\epsilon)$ and $\hat{F}_B(\epsilon)$ the empirical distribution functions of errors, relative to algorithms A and B respectively, we introduce the following

Definition 1: Algorithm A is statistically *uniformly better* than algorithm B if

$$\hat{F}_A(\epsilon) > \hat{F}_B(\epsilon), \quad \forall \epsilon \in \mathcal{E}_{AB} \quad (12)$$

with $\mathcal{E}_{AB} = \{\epsilon \in \mathbb{R} \mid 0 < \hat{F}_A(\epsilon) + \hat{F}_B(\epsilon) < 2\}$.

Set \mathcal{E}_{AB} consists of all real numbers except those for which $\hat{F}_A(\epsilon) = \hat{F}_B(\epsilon) \in \{0, 1\}$, which are not relevant for comparison. Note that when (12) holds true, $\hat{F}_A(\epsilon) \geq \hat{F}_B(\epsilon) \forall \epsilon \in \mathbb{R}$, with the inequality becoming strict for $\epsilon \in \mathcal{E}_{AB}$. In other words, algorithm A is statistically uniformly better than algorithm B when $\hat{F}_B(\epsilon)$ is upper bounded by $\hat{F}_A(\epsilon)$ for all values of ϵ except those for which $\hat{F}_A(\epsilon)$ and $\hat{F}_B(\epsilon)$ are both equal to 0 or 1. In particular, if algorithm A is statistically uniformly better than algorithm B , then it exhibits lower errors with higher probability. As a consequence, all the moments of error for algorithm A are lower than the corresponding moments for algorithm B . Moreover the median error for algorithm A , which is a more robust index of centrality [14], is also lower.

III. SIMULATION RESULTS

Figure 1 reports the empirical distribution functions of errors (10) for the detrending algorithms under analysis. The parameter λ for WQVR was chosen as the one, namely $\lambda_{\text{opt}}(i)$, that entails the minimum of (10) for each trend realization ξ_i . This choice is motivated by the need to determine the limit performance of the proposed algorithm. However, it will be shown in the next subsection that the value of λ is *not* critical.

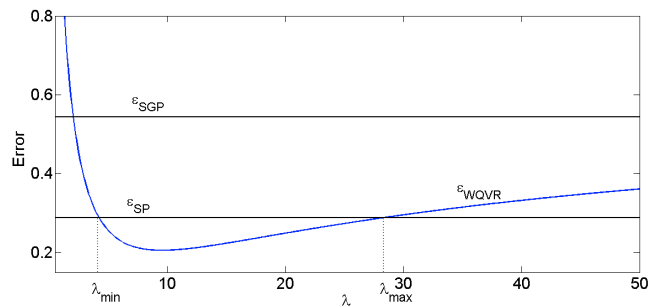


Fig. 2. Typical behavior of the error (10) for WQVR as a function of λ (blue), in comparison with the same error for SP and SGP (black).

As Figure 1 highlights, detrending by WQVR exhibits the best performance as it returns lower values of error with higher probability. In particular, it is statistically *uniformly better* than competing algorithms. Moreover, it is worth noting that the gap between the empirical distribution function of WQVR and those of competing algorithms, especially SGP, is quite high. This is a measure of the performance margin that WQVR has over the other algorithms. As regards SP, it should be remarked that it suffers from severe computational burden, which makes impossible its application to long RR series [10].

For the sake of fairness, it should be noted that being statistically uniformly better does not exclude the eventuality that for some realizations of trend SP and/or SGP could exhibit lower errors than WQVR. Nevertheless, in our simulations detrending by WQVR returned the lowest error for *any* realization of trend.

A. Robustness to variations of λ

We analyzed how sensitive the solution (8) is to variations of its controlling parameter λ . In particular, we wanted to ascertain how large the interval of values of λ is for which WQVR is still the best algorithm. The larger this interval is the more robust WQVR is to variations of λ . As shown below, such interval is *very large*, and this makes the choice of λ *not* a critical issue: a wide range of different values of λ produces similar estimates.

For a generic trend realization, the error (10) returned by WQVR as a function of λ has the typical behavior depicted in blue in Figure 2. Thus, with reference to this figure, for each trend realization ξ_i we can compute the interval $[\lambda_{\min}(i), \lambda_{\max}(i)]$ such that WQVR still yields the minimum error (10) among all algorithms, and the value $\lambda_{\text{opt}}(i) \in [\lambda_{\min}(i), \lambda_{\max}(i)]$ for which WQVR returns the minimum error (10).

Figure 3 reports in logarithmic scale the values $\lambda_{\min}(i)$, $\lambda_{\text{opt}}(i)$, and $\lambda_{\max}(i)$ (dashed lines) obtained for the 300 trend realizations and the corresponding arithmetic means $\bar{\lambda}_{\min} = 5.02$, $\bar{\lambda}_{\text{opt}} = 10.35$, $\bar{\lambda}_{\max} = 27.22$, respectively (continuous horizontal lines). As Figure 3 shows, the interval of values of λ for which WQVR returns the best performance is rather *large* for any realization. This is particularly evident when considering the average values $\bar{\lambda}_{\min}$ and $\bar{\lambda}_{\max}$, which

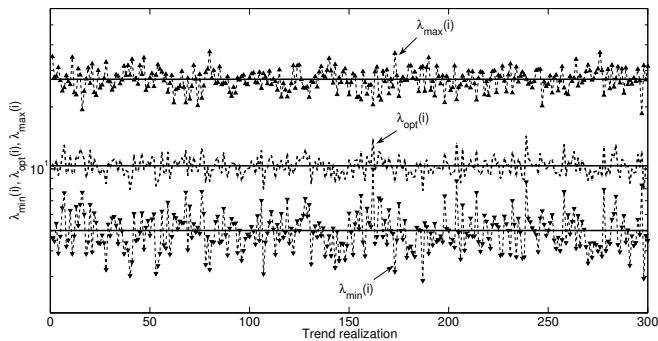


Fig. 3. $\lambda_{\min}(i)$, $\lambda_{\text{opt}}(i)$, and $\lambda_{\max}(i)$ as a function of the i th trend realization.

are approximately equidistant (in logarithmic scale) from $\bar{\lambda}_{\text{opt}}$.

In Figure 4 we report the empirical distribution functions of error (10) for WQVR choosing the same constant value of λ , respectively $\bar{\lambda}_{\min}$, $\bar{\lambda}_{\text{opt}}$, and $\bar{\lambda}_{\max}$, for all realizations of trend. Note that this is an unfavorable condition for WQVR, but for this very reason it is also a good method to assess its robustness. For ease of comparison, in Figure 4 we report the empirical distribution functions of SP and WQVR from Figure 1 as reference.

First, we compare the reference empirical distribution function of WQVR (solid line) and the corresponding function computed using $\lambda = \bar{\lambda}_{\text{opt}}$ (dash-dotted line) for all trend realizations. The two curves are almost *indistinguishable*, thus proving that even with constant λ it is possible to achieve performance nearly *identical* to the limit performance. This result is important and confirms the robustness of WQVR to variations of λ .

Second, let us consider the curves corresponding to $\lambda = \bar{\lambda}_{\max}$ and SP. The two curves are very close, thus meaning that, even choosing for λ a value larger than the double of the average optimal one (since $\bar{\lambda}_{\max}/\bar{\lambda}_{\text{opt}} = 2.6$), performance of WQVR is as good as SP. However, note that WQVR offers a great advantage in terms of computational complexity [10].

Finally, the curve corresponding to $\lambda = \bar{\lambda}_{\min}$ highlights that, even choosing a very low value of λ , i.e., half of the average optimal, the maximum error returned by WQVR is much lower than the maximum error of SP. Additionally, the median error, a more robust index of centrality, of WQVR for this choice of λ is lower than the median error of SP.

IV. CONCLUSIONS

In this work we have considered the problem of detrending for RR series. To tackle this issue, we have recently proposed a novel approach based on the notion of *weighted quadratic variation reduction*. It has been devised for unevenly sampled series, and does not rely on interpolation and resampling steps. In this paper we have extensively assessed its performance through a statistical analysis based on the empirical distribution function of the trend estimation error: the empirical distribution function provides a complete statistical description of the behavior of an algorithm over

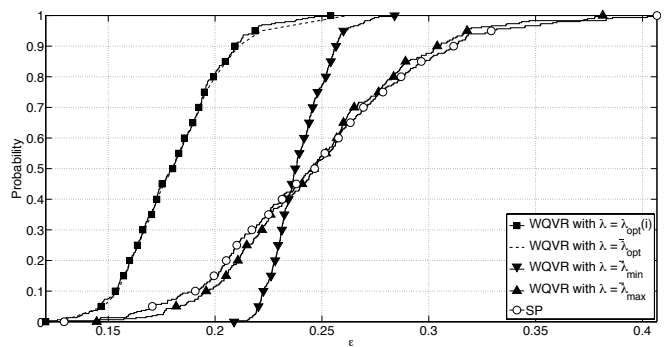


Fig. 4. Empirical distribution functions of WQVR using constant values of λ equal to $\bar{\lambda}_{\min}$, $\bar{\lambda}_{\text{opt}}$, and $\bar{\lambda}_{\max}$, and curves from Figure 1 as reference.

the whole range of estimation errors, taking account of error relative frequencies. Results confirm the effectiveness of our approach, which outperforms state-of-the-art algorithms. Further, it is *statistically uniformly better* than competing algorithms. We have shown that the algorithm is robust with respect to variations of its controlling parameter, whose value is not a critical issue. Finally, our algorithm is advantageous also in terms of computational complexity: to the best of our knowledge, it is the *fastest* algorithm for detrending RR series.

REFERENCES

- [1] Task Force of The European Society of Cardiology & The North American Society of Pacing and Electrophysiology, "Heart rate variability - Standards of measurement, physiological interpretation, and clinical use," *Eur. Heart J.*, vol. 17, pp. 354–381, 1996.
- [2] G. D. Clifford, F. Azuaje, and P. McSharry, Eds., *Advanced Methods and Tools for ECG Data Analysis*. Artech House, Inc., 2006.
- [3] L. Sörnmo and P. Laguna, *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier Academic Press, 2005.
- [4] I. Mitov, "A method for assessment and processing of biomedical signals containing trend and periodic components," *Med. Eng. Phys.*, vol. 20, pp. 660–668, 1998.
- [5] D. A. Litvack, T. F. Oberlander, and S. J. P. Carney, Laurel H., "Time and frequency domain methods for heart rate variability analysis: A methodological comparison," *Psychophysiology*, vol. 32, pp. 492–504, 1995.
- [6] M. Rantavain, P. Ranta-aho, and P. Karjalainen, "An advanced detrending method with application to HRV analysis," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 2, pp. 172–175, feb. 2002.
- [7] G. D. Clifford and L. Tarassenko, "Quantifying errors in spectral estimates of HRV due to beat replacement and resampling," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 630–638, 2005.
- [8] G. B. Moody, "Spectral analysis of heart rate without resampling," *Computers in Cardiology*, vol. 20, pp. 715–718, 1993.
- [9] A. Eleuteri, A. C. Fisher, D. Groves, and C. J. Dewhurst, "An efficient time-varying filter for detrending and bandwidth limiting the heart rate variability tachogram without resampling: MATLAB open-source code and internet web-based implementation," *Comp. Math. Meth. Med.*, 2012.
- [10] V. Villani and A. Fasano, "Fast detrending of unevenly sampled series with application to HRV," *IEEE Comput. Cardiol.*, vol. 40, pp. 417–420, 2013.
- [11] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, "A dynamical model for generating synthetic electrocardiogram signals," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 289–294, 2003.
- [12] A. Gut, *Probability: A Graduate Course*. Springer, 2013.
- [13] O. Thas, *Comparing Distributions*, 1st ed. Springer Publishing Company, Incorporated, 2009.
- [14] P. J. Huber and E. M. Ronchetti, *Robust Statistics*. Wiley, 2009.