Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: http://ees.elsevier.com

Bayesian source separation of electrical bioimpedance signals

Christof Pichler [a,b], Sascha Ranftl [a,b,*], Arnulf Heller [c], Enrico Arrigoni [a], Wolfgang von der Linden [a,b]

[a] Institute of Theoretical and Computational Physics, Graz University of Technology, Petersgasse 16, 8010 Graz, Austria
[b] Graz Center of Computational Engineering, Graz University of Technology, Krenngasse 37/I, 8010 Graz, Austria
[c] ImPress MedTech GmbH, ZWT, Neue Stiftingtalstraße 2A, 8010 Graz, Austria

ARTICLE INFO

Article history:
Received 17 August 2020
Received in revised form 22 January 2021
Accepted 27 February 2021
Available online xxx

Keywords
Uncertainty quantification
Bayesian probability theory
Source separation
Electrical bioimpedance
Impedance cardiography

ABSTRACT

For physicians, it is often crucial to monitor hemodynamic parameters to provide appropriate treatment for patients. Such hemodynamic parameters can be estimated via electrical bioimpedance (EBI) signal measurements. Time dependent changes of the measured EBI signal occur due to several different phenomena in the human body. Most of the time one is just interested in a single component of the EBI signal, such as the part caused by cardiac activities, wherefore it is necessary to decompose the EBI signal into its different source terms. The changes of the signal are mostly caused by respiration and cardiac activity (pulse). Since these fluctuations are periodic in sufficiently small time windows, the signal can be approximated by a harmonic series with two different fundamental frequencies and an unknown number of higher harmonics. In this work, we present Bayesian robability heory as the adequate and rigorous method for this decomposition. The proposed method allows, in contrast to other methods, to consistently identify the model-function, compute parameter estimates and predictions, and to quantify uncertainties. Further, the method can handle a very low signal-to-noise ratio. The results suggest that EBI-based estimation of hemodynamic parameters and their monitoring can be improved and its reliability assessed.

1. Introduction

There are many experimental methods to determine hemodynamic parameters. Most of them are invasive and therefore, electrical bioimpedance (EBI) measurements, pulse transit time methods [23] or photoplethysmography [7] provide interesting non-invasive alternatives [24]. The EBI method is well established, cheap and relatively easy-to-use. Moreover, it is suitable for time-continuous and long term measurements. Almost all periodic changes in the EBI signal are caused by respiration and cardiac activity [21]. Blood has a lower resistivity than other parts of the human body, which causes the injected electric current to flow mainly through blood vessels. Thus, local changes of blood distribution due to cardiac activity cause periodic changes in the EBI signal for sufficiently small time windows. Moreover, the changing air volume in the lungs during the breathing cycles also causes periodic changes in the resistivity and therefore in the EBI signal [4]. Since the EBI signal of a living human body mostly depends on its cardiovascular activity [8], it provides a suitable tool for estimating some hemodynamic parameters. For that purpose, however, it is crucial to separate the signal into its source terms. A few examples of medical applications, where the decomposition of the signal improves the results, are heart rate detection [6], estimation of the cardiac output [10] or the reconstruction of the central arterial pressure waveform [19]. Several techniques addressing EBI signal separation exist in the literature, such as independent component analysis [22], ensemble averaging [9] or adaptive filtering [25]. Krivoshei et al. [18] proposed an adaptive phase-locked loop method and Butsenko et al. [5] proposed to separate the signal via sparse reconstruction. However, none of these methods are able to quantify also the uncertainties.

In contrast, Bayesian probability theory allows not only to decompose the signal, but also to determine the uncertainties of both the decomposed signal and of the estimated parameters. Another major advantage of the Bayesian approach is the capability to handle measurements with a very low signal to noise ratio (SNR), as will be shown in this article. Our method could possibly be used to improve estimates and uncertainties of e.g. hemodynamic parameters that are based on decomposed EBI signals.

The remainder of this paper is organized as follows. In Section 2, we outline the basic principles of Bayesian probability theory for the readers' convenience. In Section 3, we develop the method and in Section 4 we validate the method with artificial data. In Section 5 we illustrate an application to real EBI measurements.

2. Bayesian probability theory

Here, we face an inverse problem, in which the question is how to find the causes given observed data and some prior knowledge. Bayesian probability theory is a consistent and successful tool to deal

* Corresponding author at: Institute of Theoretical and Computational Physics, Graz University of Technology, Petersgasse 16, 8010 Graz, Austria.
E-mail address: ranftl@tugraz.at (S. Ranftl)

https://doi.org/10.1016/j.bspc.2021.102541
1746-8094/

with such problems. Here, we want to outline the main aspects of this approach. For more detailed information about Bayesian probability theory we refer to [14,26,20].

2.1. Probability

Probability is a measure for the correctness of a proposition (hypothesis), corresponding to a statement that can only be true or false. Probability theory can therefore be viewed as the generalization of the propositional calculus to the case of partial truth. Probabilities can have values between 0 and 1, corresponding to the impossible and to the certain event, respectively. It is important to distinguish between the probability $p(x)$ that a certain proposition x is true without further knowledge and the *conditional probability* $p(x \mid y)$ that the same proposition holds conditioned on some given information y, e.g. some measured data. Like in Boolean algebra, propositions can be combined by the logical AND ($\wedge$) and a logical OR ($\vee$). In Bayesian probability theory the probability $p(x \wedge y)$ is denoted by $p(x, y)$ for better readability. Probabilities have to fulfill two basic rules: the sum rule and the product rule [14,26,20]. With these rules, it is possible to derive the marginalization rule and Bayes' theorem.

2.2. Marginalization rule

The marginalization rule for discrete variables (propositions) reads

$$p(x \mid I) = \sum_{i=1}^{M} p(x \mid y_i, I)\, p(y_i \mid I), \tag{1}$$

where $\{y_i\}_M$ is a mutually exclusive and exhaustive set of M propositions to fulfill the normalization requirement [26].

The proposition I contains all the relevant background information. For continuous variables, the marginalization rule takes an integral form

$$p(x \mid I) = \int p(x \mid y, I) p(y \mid I) \mathrm{d}y. \tag{2}$$

We use the same symbol $p(\bullet)$ for probabilities and probability density functions (pdf). Which one we are referring to will be clear depending on whether the relevant variable is discrete or continuous.

2.3. Bayes' theorem

Bayes' theorem follows directly from the product rule [14,26,20] and reads

$$p(x \mid y, I) = \frac{p(y \mid x, I) \cdot p(x \mid I)}{p(y \mid I)}. \tag{3}$$

Bayes' theorem is valid for both cases of discrete or continuous variables x and y. Accordingly, we have to use the corresponding probabilities or pdf. Given, e.g. a set of parameters x and measured data y, then Bayes' theorem (3) allows to determine the probability distribution $p(x \mid y, I)$ for the parameters given the data. To achieve this, one needs knowledge of the likelihood $p(y \mid x, I)$, that encodes the error statistics of the experiment, which in many cases is described by a Gaussian. The prior probability $p(x \mid I)$ represents the state of knowledge or ignorance about the parameters *before* the experimental data are known. The denominator $p(y \mid I)$ in (3) is called the evidence (5).

We now present a brief discussion how these ingredients can be determined in a given physical problem. For a detailed presentation, we refer to the ample literature.

2.3.1. Prior probability

There are several principles to derive a prior distribution [28], such as transformation invariance [13,16], the maximum entropy principle [11,12,15] or reference priors [2]. A discussion of these approaches is beyond the scope of the present paper and we refer to the available literature. Here, we will merely present some models useful for the specific problem.

2.3.2. Likelihood function

The likelihood is the probability (density) for the data y given all parameters and the underlying physical model. More specifically, if we assume a certain model to describe the data and we know its parameters, we should be able to determine the ideal values of the data $\hat{y}$. Since experimental data always have noise, their measured value will be

$$y = \hat{y} + \eta, \tag{4}$$

where η is the experimental noise/uncertainty, which in many cases is described by a Gaussian distribution. The term "likelihood" instead of probability is used when referring to variables in the conditional complex behind the conditional mark |, i.e. for the parameters x the probability (pdf) $p(y|x)$ is considered as likelihood. If $p(y|x_1) > p(y|x_2)$, it is reasonable to assume that x_1 is more likely to be the correct parameter set describing the data than x_2. But what we really need to reach this conclusion is $p(x_1|y) > p(x_2|y)$ [26,20].

If the distribution of the noise is known, one can easily write down the likelihood function of the measured data. This will be discussed in more detail in chapter 3.

2.3.3. Posterior probability

The posterior probability gives the probability that, knowing the data, the model is described by certain values of parameters.

2.3.4. Evidence

As one can see from (3), in the context of parameter estimation, the evidence merely gives a constant term that can be computed at the end of the calculation. However, the evidence is important when one has to select between different models, as will be discussed in Section 4.1 [26,20]. The evidence is calculated by marginalization (2), which in the case of continuous variables x and y becomes

$$p(y \mid I) = \int p(y \mid x, I)\, p(x \mid I) \mathrm{d}x. \tag{5}$$

2.4. Moments of probability distributions

Information about a pdf is given by their momenta. The nth moment is given by

$$\langle x^n \rangle = \int x^n p(x \mid I) \mathrm{d}x. \tag{6}$$

The first moment ($n = 1$) is also called the mean. The variance is defined as

$$\mathrm{Var}(x) = \left\langle x^2 \right\rangle - \langle x \rangle^2. \tag{7}$$

The standard deviation is defined as the square root of the variance,

$$\mathrm{Std}(x) = \sqrt{\mathrm{Var}(x)}, \tag{8}$$

and is a measure of the uncertainty of the estimated mean.

3. Bayesian data analysis of the bioimpedance signal

3.1. Model function

As the changes of the EBI signal are assumed to be periodic in time, we approximate this by the lowest order terms of a harmonic series expansion. In principle, the EBI signal is determined by the respiration and the cardiac activity [21]. Therefore, we assume that there are two fundamental frequencies in the model function. Due to the wave form of the signal, there are also integer multiples of these frequencies (higher harmonics) present in the spectrum [17]. This is confirmed by the frequency spectrum of the experimental EBI signals showing dominant peaks that can be attributed to two fundamental frequencies and their integer multiples. The model function should therefore consist of an harmonic series that depends on two fundamental frequencies ω_r and ω_p including their n_r and n_p higher harmonics, where r denotes respiration and p denotes pulse.

$$f(t) = \sum_{k=1}^{\frac{m}{2}} \left(B_k \ \sin(\omega_k \ t) + B_{k+\frac{m}{2}} \ \cos(\omega_k \ t) \right) \tag{9}$$

with $m = 2(n_r + n_p)$ and

$$\omega_k = \begin{cases} k \cdot \omega_r : & k \le n_r \\ (k - n_r) \cdot \omega_p : & k > n_r \end{cases}$$

For better readability we write the function $f(t)$ in a more compact way

$$f(t) = \sum_{j=1}^{m} B_j \ G_j(t). \tag{10}$$

The expansion order m is an even integer, as we always have the pair of cos and sin. The basis functions $G_j(t)$ are extracted by comparison with (9). For simplicity, we restrict the analysis to the cases $n_p = n_r$ or $n_p = n_r + 1$. In other words, if $m/2$ is even $n_p = n_r$, otherwise we use $n_p = n_r + 1$. The case $n_r \approx n_p$ turned out to be where we found the highest posterior probability in most cases we tested. That is why we used this restriction to keep the computational effort as small as possible.

3.2. Bayesian spectrum analysis and parameter estimation

Bayesian inference allows to estimate the optimal values of model parameters, as well as their uncertainty. In our case these parameters are B_j and ω_α with $\alpha = \{r, p\}$ in (9) and the experimental noise level σ (see below). This idea is outlined in the book [3]. The EBI measurements provide a set of data d_i at times t_i with $i = 1, \ldots, N$. We assume that each measured data point d_i is equal to the model $f(t_i)$ corrupted by additive noise η_i,

$$d_i = f(t_i) + \eta_i, \tag{11}$$

where η_i is Gaussian distributed, i.e. $\eta_i \sim \mathcal{N}(0, \sigma^2)$. For simplicity, the experimental noise level σ is assumed to be the same for every data point and uncorrelated at different time instances. Therefore, for a given value of the parameters B_j, ω_α, σ, the likelihood function, i.e. the probability to measure a given set of values d_i for the data, is given by

$$\prod_{i=1}^{N} p\left(d_i \mid B_j, \omega_\alpha, \sigma, m, I\right) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{\Phi}{2\sigma^2}}, \tag{12}$$

with

$$\Phi = \sum_{i=1}^{N} \left(d_i - f(t_i)\right)^2. \tag{13}$$

When expanding the square, Φ contains mixing terms of the form $B_j B_{j'}$ which are difficult to treat. As suggested in [3], it is convenient to replace the functions $G_j(t)$ with an orthonormal set of functions $H_k(t)$ associated to new amplitudes A_k. This procedure can be better described within a matrix-vector notation whereby $\boldsymbol{d}$, $\boldsymbol{f}$, $\boldsymbol{B}$, and $\boldsymbol{A}$ are vectors with components d_i, $f(t_i)$, B_j, and A_j, and G and H are matrices with elements

$$G_{ik} = G_k(t_i), \quad \text{and} \quad H_{ik} = H_k(t_i).$$

The orthogonalization is achieved by the transformation

$$H = GS^{-\frac{1}{2}} \tag{14}$$

with S being the overlap matrix of the original basis functions, i.e.

$$S = G^T G,$$

leading to the orthonormality condition

$$H^T H = \mathbb{1} \tag{15}$$

The coefficients transform as

$$\boldsymbol{A} = S^{\frac{1}{2}} \boldsymbol{B}, \tag{16}$$

so that (10) can be written as

$$\boldsymbol{f} = G\boldsymbol{B} = H\boldsymbol{A}. \tag{17}$$

In the new orthonormal basis the argument (13) of the exponential in the likelihood simplifies to

$$\Phi = \hat{\Phi} + (\boldsymbol{A} - \boldsymbol{h})^T (\boldsymbol{A} - \boldsymbol{h}), \tag{18}$$

$$\text{with} \qquad \boldsymbol{h} = \boldsymbol{d}\, H, \tag{19}$$

$$\hat{\Phi} = \boldsymbol{d}^2 - \boldsymbol{h}^2. \tag{20}$$

We have now an expression of the likelihood (12) in terms of the new coefficients $\boldsymbol{A}$

$$p(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, m, I) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{\Phi}{2\sigma^2}}, \tag{21}$$

with Φ given in (18). To estimate the optimal values of the model parameters we need the pdf $p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid \boldsymbol{d}, m, I)$ of the amplitudes $\boldsymbol{A}$, frequencies $\boldsymbol{\omega} = \{\omega_p, \omega_r\}$, and noise level σ given the information on the measured data $\boldsymbol{d}$ and the expansion order m. The set of times $\boldsymbol{t}$, at which the signal values $\boldsymbol{d}$ are measured, are put into the background information, here represented by I, as it is not explicitly addressed in the calculations. Bayes' theorem allows to write

$$p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid \boldsymbol{d}, m, I) = \frac{p(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, m, I)\ p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid m, I)}{p(\boldsymbol{d} \mid m, I)}. \tag{22}$$

Next, we will discuss the priors for the parameters. First of all, these parameters are logically independent and therefore the pdf factorizes

$$p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid m, I) = p(\boldsymbol{A} \mid m, I) p(\boldsymbol{\omega} \mid I) p(\sigma \mid I).$$

We have omitted the conditioning on m in the last two terms as $\boldsymbol{\omega} = \{\omega_r, \omega_p\}$ and σ are independent of the number of higher harmonics. For all cases, we assume little prior knowledge and, therefore, use un-informative priors. We begin with the noise level σ. We assume that we know nothing about σ except that it is a scale parameter, i.e. it lies in $\mathbb{R}_+ = (0, \infty)$ and the pdf should be invariant against scaling $\sigma \to a\sigma$ with an arbitrary factor a, otherwise it would make a difference which units we use. Then, Jeffreys' prior (see e.g. [20]) has to be used

$$p(\sigma \mid I) = 1/\sigma. \tag{23}$$

The fact that we use an almost completely ignorant prior is reflected in the fact that the latter is not normalizable. This is no problem as long

as the final result can be normalized, which is generally the case if the data entering via the likelihood is meaningful.

In the case of the two fundamental frequencies ω_α with $\alpha \in \{r, p\}$ we assume more knowledge than in the case of the noise level, because we know roughly the physiological range. Let us denote these intervals as $\Omega_\alpha = (\Omega_\alpha^l, \Omega_\alpha^u)$. Then the uniform prior is constant within this interval and zero outside of it,

$$p(\omega_\alpha) = \frac{1}{\Omega_\alpha^u - \Omega_\alpha^l}\, \theta\left(\Omega_\alpha^u < \omega_\alpha < \Omega_\alpha^u\right). \tag{24}$$

We have chosen reasonable values $\Omega_r = (0.6,\ 6.0)$ rad/s and $\Omega_p = (13.5,\ 17.5)$ rad/s. These intervals take into account that the natural variation of the respiratory part is greater than that of the pulsatile part. The exact choice of these intervals is generally quite unimportant for the final result, as it is dominated by data constraints.

For the amplitudes we could also use uniform priors, but instead it is advantageous to use the so-called signal-power prior, proposed by [3]

$$p(\boldsymbol{A}|\delta) \propto \exp\left(-\frac{1}{2\delta^2}\sum_i (f(t_i))^2\right).$$

The $\boldsymbol{A}$-dependence enters via the model functions, defined in (17). The properly normalized prior then reads

$$p(\boldsymbol{A}|m, \delta, I) = (2\pi)^{-\frac{m}{2}}\delta^{-m} e^{-\frac{1}{2\delta^2}\boldsymbol{A}^2}, \tag{25}$$

where we have used the orthonormality of the basis functions. To get rid of the unknown hyper-parameter δ we invoke again the marginalization rule

$$p(\boldsymbol{A}|m, I) = \int d\delta\, p(\boldsymbol{A}|m, \delta, I)\ p(\delta|I). \tag{26}$$

Since δ is a scale parameter we use again Jeffreys' prior, i.e. $p(\delta|I) \propto 1/\delta$, and obtain

$$p(\boldsymbol{A}|m, I) = \frac{1}{\mathscr{Z}}(2\pi)^{-\frac{m}{2}}\left(\frac{\boldsymbol{A}^2}{2}\right)^{-\frac{m}{2}}\Gamma\left(\frac{m}{2}\right), \tag{27}$$

with $\Gamma(\cdot)$ being the Gamma-function and $\mathscr{Z}$ the normalization constant. As Jeffreys' prior is not normalizable, however, the resulting prior in (27) also loses its normalizability, but, as in the case of Jeffreys' prior, this is irrelevant for our purposes, as in combination with the likelihood function the final results will be normalizable again.

Good data typically lead to likelihood functions which are sharply peaked as function of the parameters, whereas the prior typically varies little within the parameter range of the peak. Therefore, the $\boldsymbol{A}$-dependence of the prior has negligible influence on the parameter estimation. On the other hand, its m-dependence will become important for model comparison (Section 3.4). This is due to the following property. The likelihood increases monotonically with increasing m, as the misfit decreases. On the other hand, the prior typically decreases with increasing m. Here, the prior cannot be ignored and only the interplay of both terms leads to a peak in the probability for m.

Now that we have defined the priors, we can calculate all the marginal probabilities for the parameters by the marginalization rule (2). We begin with the probability for the frequencies $\boldsymbol{\omega}$ which is obtained by integrating with respect to all other parameters

$$\begin{aligned} p(\boldsymbol{\omega} \mid \boldsymbol{d}, m, I) &= \int d\sigma \int dV_{\boldsymbol{A}}\, p(\boldsymbol{\omega}, \boldsymbol{A}, \sigma \mid \boldsymbol{d}, m, I) \\ &= p(\boldsymbol{\omega} \mid I)\ \int d\sigma\, p(\sigma \mid I) \\ &\quad \times \int dV_{\boldsymbol{A}}\, p(\boldsymbol{d} \mid \boldsymbol{\omega}, \boldsymbol{A}, \sigma, m, I)\ \ p(\boldsymbol{A} \mid m, I). \end{aligned}$$

Here the infinitesimal integration measure is given by $\mathrm{d}V_{\boldsymbol{A}} = \prod_i dA_i$. Due to the large number of data points, the Gaussian likelihood is sharply peaked in $\boldsymbol{A}$ as compared to the slow variation of the prior $p(\boldsymbol{A} \mid m, I)$, see (27), so we can replace the prior by its value at the maximum of the likelihood $\boldsymbol{A} \to \boldsymbol{h}$ (see (18)). Then all integrals can be performed analytically, resulting in

$$p(\boldsymbol{\omega} \mid \boldsymbol{d}, m, I) = \frac{1}{\mathscr{Z}}\hat{\Phi}^{\frac{m-N}{2}}. \tag{28}$$

Here, N is the number of data points and m is the number of basis functions, and $\hat{\Phi}$ was defined in (18).

For the other quantities of interest we have to integrate over the frequencies. As they enter non-linearly in the basis functions and corresponding orthogonalization, an analytic integration is impossible. In this case, it is expedient to first integrate over the amplitudes and the noise level, which can be done analytically. That leads to probability densities *conditioned* on the frequencies $\boldsymbol{\omega}$. Finally, these expressions have to be integrated numerically over frequencies $\boldsymbol{\omega}$ weighted by the pdf $p(\boldsymbol{\omega} \mid \boldsymbol{d}, m, I)$, given in (28).

In these steps, the pdf $p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, m, I)$ is a crucial quantity. It can be computed analytically in a similar way as $p(\boldsymbol{\omega} \mid \boldsymbol{d}, m, I)$ resulting in

$$p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, m, I) = \frac{1}{\mathscr{Z}}\sigma^{-(N-m+1)} e^{-\frac{\hat{\Phi}}{2\sigma^2}}. \tag{29}$$

From this one obtains the analytic expression for the νth moment (6) of the noise level conditioned on $\boldsymbol{\omega}$:

$$\langle\sigma^\nu\rangle_{\boldsymbol{\omega}} = \frac{\Gamma\left(\frac{N-m-\nu}{2}\right)}{\Gamma\left(\frac{N-m}{2}\right)}\cdot\left(\frac{\hat{\Phi}}{2}\right)^{\frac{\nu}{2}} \tag{30}$$

and reads off the mean value of σ^2

$$\left\langle\sigma^2\right\rangle_{\boldsymbol{\omega}} = \frac{\hat{\Phi}}{N-m-2}. \tag{31}$$

According to (20), $\hat{\Phi}$ is the misfit, and $\langle\sigma^2\rangle_{\boldsymbol{\omega}}$ is therefore a reasonable estimate for σ^2, which is the squared noise level of the data points. For the variance of σ^2 we obtain

$$\left\langle\left(\Delta\sigma^2\right)^2\right\rangle_{\boldsymbol{\omega}} = \frac{2\left\langle\sigma^2\right\rangle_{\boldsymbol{\omega}}^2}{N-m-4}. \tag{32}$$

For $N \gg m$, Eq. (32) shows that the variance/uncertainty of the squared noise level is $O(1/N)$ and therefore very small due to the large number of data points supporting the estimate of the few unknown parameters. The same is true for the noise level σ itself, which can therefore be approximated by a Dirac distribution

$$p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, m, I) \approx \delta(\sigma - \hat{\sigma}_\omega), \tag{33}$$

$$\hat{\sigma}_\omega = \sqrt{\left\langle\sigma^2\right\rangle_\omega}. \tag{34}$$

Next we compute the conditional pdf for the amplitudes

$$p(\boldsymbol{A} \mid \boldsymbol{\omega}, \boldsymbol{d}, m, I) \quad = \int d\sigma\ \ p(\boldsymbol{A} \mid \boldsymbol{\omega}, \sigma, \boldsymbol{d}, m, I)\cdot p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, m, I). \tag{35}$$

With this, together with (33), we can determine the mean and the covariance of the amplitudes conditioned on $\boldsymbol{\omega}$ (details are in Appendix A):

$$\langle A_k\rangle_{\boldsymbol{\omega}} = h_k, \tag{36}$$

$$\langle\Delta A_k \Delta A_{k'}\rangle_{\boldsymbol{\omega}} = \hat{\sigma}_{\boldsymbol{\omega}}^2\delta_{\mathrm{kk'}}. \tag{37}$$

As mentioned before, these estimates have to be integrated numerically over frequencies $\boldsymbol{\omega}$ weighted by the pdf $p(\boldsymbol{\omega} \mid \boldsymbol{d}, m, I)$, given in (28).

3.3. Signal estimation and uncertainty band

The aim of this section is to estimate the signal and calculate the associated uncertainty band. In the following equations, f_i represents the value of the model function at time t_i for a given set m of higher harmonics. For better readability we do not write m explicitly in the conditional part of the pdfs. The starting point is the pdf for the value f_i of the signal [20, p. 351].

$$\begin{aligned} p(f_i \mid \boldsymbol{d}, I) &= \int \mathrm{dV}_{\boldsymbol{\omega}} \int \mathrm{dV}_{\boldsymbol{A}} \int d\sigma\, p(f_i \mid \boldsymbol{A}, \sigma, \boldsymbol{\omega}, \boldsymbol{d}, I) \\ &\quad \times p(\boldsymbol{A}, \sigma, \boldsymbol{\omega} \mid \boldsymbol{d}, I). \end{aligned} \tag{38}$$

Here the infinitesimal integration measure is given by $\mathrm{dV}_{\boldsymbol{\omega}} = \prod_i d\omega_i$. For given parameters $\boldsymbol{\omega}$ and $\boldsymbol{A}$, the function value is exactly defined by (17), i.e.

$$\boldsymbol{f} = H\boldsymbol{A}. \tag{39}$$

In this case, the probability $p(f_i \mid \boldsymbol{A}, \sigma, \boldsymbol{\omega}, \boldsymbol{d}, I)$ is a delta distribution $\delta(f_i - (H\boldsymbol{A})_i)$, where $(H\boldsymbol{A})_i$ denotes the ith component of the vector $H\boldsymbol{A}$. This leads in combination with (38) and (6) to the following expression for the expectation value of the νth moment of the signal f_i at time t_i

$$\begin{aligned} \langle f_i^{\nu} \rangle &= \int \mathrm{df}_i f_i^{\nu} p(f_i \mid \boldsymbol{d}, \sigma, I) \\ &= \iiint \mathrm{dV}_{\boldsymbol{\omega}} \mathrm{dV}_{\boldsymbol{A}} d\sigma (H\boldsymbol{A})_i^{\nu} p(\boldsymbol{A}, \sigma, \boldsymbol{\omega} \mid \boldsymbol{d}, I) \\ &= \int \mathrm{dV}_{\boldsymbol{\omega}} p(\boldsymbol{\omega} \mid \boldsymbol{d}, I) \int \mathrm{dV}_{\boldsymbol{A}} (H\boldsymbol{A})_i^{\nu} \\ &\quad \times \int d\sigma p(\boldsymbol{A} \mid \sigma, \boldsymbol{\omega}, \boldsymbol{d}, I) p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, I). \end{aligned} \tag{40}$$

As explained above, the pdf for σ can very reliably be approximated by a Dirac distribution (33). Thus, (40) simplifies to

$$\langle f_i^{\nu} \rangle = \int \mathrm{dV}_{\boldsymbol{\omega}}\, p(\boldsymbol{\omega} \mid \boldsymbol{d}, I)\ \langle (H\boldsymbol{A})_i^{\nu} \rangle_{\boldsymbol{\omega}}, \tag{41}$$

with the definition of the conditional mean as before

$$\langle (H\boldsymbol{A})_i^{\nu} \rangle_{\boldsymbol{\omega}} = \int \mathrm{dV}_{\boldsymbol{A}}\ (H\boldsymbol{A})_i^{\nu}\, p(\boldsymbol{A} \mid \hat{\sigma}_{\omega}, \boldsymbol{\omega}, \boldsymbol{d}, I). \tag{42}$$

The pdf $p(\boldsymbol{A} \mid \hat{\sigma}_{\omega}, \boldsymbol{\omega}, \boldsymbol{d}, I)$ as derived in Appendix A is given by

$$p(\boldsymbol{A} \mid \hat{\sigma}_{\omega}, \boldsymbol{\omega}, \boldsymbol{d}, I) = \hat{\sigma}_{\omega}^{-m} (2\pi)^{-\frac{m}{2}}\, e^{-\frac{1}{2\hat{\sigma}_{\omega}^2}(\boldsymbol{A}-\boldsymbol{h})^2}. \tag{43}$$

For the first moment we have

$$\langle (H\boldsymbol{A})_i \rangle_{\boldsymbol{\omega}} = (H\boldsymbol{h})_i. \tag{44}$$

Then the integral in (41) with respect to the amplitudes yields

$$\langle f_i \rangle = \int \mathrm{dV}_{\boldsymbol{\omega}}\, p(\boldsymbol{\omega} \mid \boldsymbol{d}, I)\, (H\boldsymbol{h})_i. \tag{45}$$

It should be noted that, by definition, H and $\boldsymbol{h}$ depend on $\boldsymbol{\omega}$. The squared uncertainty, that shall be denoted by $(\Delta f_i)^2$, is defined as

$$(\Delta f_i)^2 = \langle f_i^2 \rangle - \langle f_i \rangle^2. \tag{46}$$

The evaluation, outlined in Appendix B, yields

$$\begin{aligned} (\Delta f_i)^2 &= \int \mathrm{dV}_{\boldsymbol{\omega}}\, p(\boldsymbol{\omega} \mid \boldsymbol{d}, I) \\ &\quad \times \left[\left\langle \left[(H\boldsymbol{h})_i - \langle (H\boldsymbol{h})_i \rangle_{\boldsymbol{\omega}} \right]^2 \right\rangle_{\boldsymbol{\omega}} + (\mathrm{HH}^T)_{\mathrm{ii}} \hat{\sigma}_{\boldsymbol{\omega}}^2 \right]. \end{aligned} \tag{47}$$

The remaining integrals in (45) and (47) over the frequencies $\boldsymbol{\omega}$, based on the pdf for $\boldsymbol{\omega}$ given in (28), are performed numerically by the nested sampling technique [27], which turned out to be more efficient than standard numerical integration schemes. In this work, the analytic evaluation of the integral over σ was possible because $N \gg m$, whereas the analytic integration over $\boldsymbol{A}$ relied on a slowly varying prior for $\boldsymbol{A}$. If these conditions are not met, nested sampling can also be used for numerical integration over these parameters.

We now additionally discuss the special case that $p(\boldsymbol{\omega} \mid \boldsymbol{d}, I)$ is sharply peaked. Then $p(\boldsymbol{\omega} \mid \boldsymbol{d}, I)$ can be approximated by a delta-distribution, centered at $\hat{\omega}$, and the first contribution to the uncertainty vanishes and we have

$$(\Delta f_i)^2 \approx (\mathrm{HH}^T)_{\mathrm{ii}} \hat{\sigma}_{\hat{\boldsymbol{\omega}}}^2, \tag{48}$$

where H now depends on $\hat{\boldsymbol{\omega}}$. Hence, the first contribution in (47) stems from the uncertainty in $\boldsymbol{\omega}$ and the second is due to the uncertainty of the amplitudes, which is proportional to $\hat{\sigma}_{\hat{\boldsymbol{\omega}}}$. This statement needs clarification: The uncertainty of $\boldsymbol{\omega}$ is given by the pdf $p(\boldsymbol{\omega} \mid \boldsymbol{d}, I)$ in (28). Still for the case of negligible uncertainty in $\boldsymbol{\omega}$ we can obtain a rough estimate of the size of the signal uncertainty. To this end, we consider the mean square signal

$$\sum_i \langle f_i \rangle^2 = \boldsymbol{h}^T \underbrace{H^T H}_{=1\ 1} \boldsymbol{h} = \boldsymbol{d}^T H H^T \boldsymbol{d} = \boldsymbol{d}^T \mathscr{P} \boldsymbol{d}.$$

$\mathscr{P}$ is a projection operator into the subspace spanned by the model functions. For a rough estimate we can assume $\mathscr{P} \approx 1\ 1$, which means that the misfit is small. Then

$$\sum_i \langle f_i \rangle^2 \approx \boldsymbol{d}^T \boldsymbol{d} = N\overline{d^2}$$

and hence

$$\sum_i (\Delta f_i)^2 = \mathrm{tr}\left(HH^T\right) \hat{\sigma}_{\hat{\boldsymbol{\omega}}}^2 = \mathrm{tr}\left(H^T H\right) \hat{\sigma}_{\hat{\boldsymbol{\omega}}}^2 = m\hat{\sigma}_{\hat{\boldsymbol{\omega}}}^2.$$

The relative uncertainty is then given by the ratio

$$\varepsilon = \sqrt{\frac{\sum_i (\Delta f_i)^2}{\sum_i \langle f_i \rangle^2}} = \sqrt{\frac{m}{N}} \sqrt{\frac{\hat{\sigma}_{\hat{\boldsymbol{\omega}}}^2}{\overline{d^2}}}.$$

The first factor is the inverse square root of the number of data points per unknown amplitude and the second factor is the noise-to-signal-ratio. This is a very reasonable behavior.

As discussed above, we are interested in separating the total signal into the respiratory and pulsatile parts. This is obtained in a very similar way. Let $\boldsymbol{f}^{\alpha}$ $\alpha \in \{r, p\}$ be the two components of the signal. These are defined as a sum over a particular subset of indices j in (10). In the matrix/vector notation of (17) we obtain $\boldsymbol{f}^{\alpha}$ by setting to zero the columns of the matrix G corresponding to the parts of the signal we are not considering. This can most elegantly be achieved by introducing the diagonal matrix P^{α} with diagonal components

$$p_{\mathrm{jj}}^{\alpha} = \begin{cases} 1 & G_j \text{is part of component } \alpha \text{ of the signal} \\ 0 & \text{otherwise} \end{cases} \tag{49}$$

Then we have

$$\boldsymbol{f}^{\alpha} = GP^{\alpha}\boldsymbol{B}.$$

Next we use the transformation to the orthonormal basis functions in (14) and (16)

$$\boldsymbol{f}^{\alpha} = \underbrace{HS^{\frac{1}{2}} P^{\alpha} S^{-\frac{1}{2}}}_{\equiv H^{\alpha}} \boldsymbol{A} = H^{\alpha}\boldsymbol{A}. \tag{50}$$

Hence, for the mean value or the uncertainty of f_i^{α} we merely have to replace H by H^{α} in (45) and (47).

3.4. Model comparison

The expansion order m, describing the number of higher harmonics in the model function (10), is still unknown. Different m can be viewed as different models with different complexity and our goal in this section is to determine the probability for a model given the data. This can be achieved by Bayes' theorem.

$$p(m \mid \boldsymbol{d}, I) = \frac{p(\boldsymbol{d} \mid m, I) p(m \mid I)}{p(\boldsymbol{d} \mid I)}. \tag{51}$$

Since we do not prefer one model over another, the prior $p(m \mid I)$ has the same value for every model. The evidence $p(\boldsymbol{d} \mid I)$ is also a constant as far as the dependence on m is concerned and, therefore, we just have to compute the likelihood $p(\boldsymbol{d} \mid m, I)$ and fix the normalization afterwards.

The pdf $p(\boldsymbol{d} \mid m, I)$ is actually a marginal likelihood, since in comparison to the full likelihood $p(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, m, I)$ in (21) the parameters are missing. The marginal likelihood is obtained by the marginalization rule

$$\begin{aligned} p\big(\boldsymbol{d} \mid m, I\big) &= \int \mathrm{dV}_{\omega}\; p(\boldsymbol{\omega} \mid m, I) \int d\sigma\;\; p(\sigma \mid I) \\ &\quad \times \int \mathrm{dV}_{\boldsymbol{A}}\; p\big(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, m, I\big)\, p\big(\boldsymbol{A} \mid m, I\big). \end{aligned} \tag{52}$$

In contrast to the case of parameter estimation, the un-informative prior for the amplitudes (25) now plays a crucial role. Due to its m-dependence it penalizes models with larger m, i.e. more complex models (so-called Ockham factor). The likelihood function $p(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, m, I)$ is a Gaussian (21), which is sharply peaked as a function of the amplitudes as compared to the weak variation of the prior. Therefore, we can replace $\boldsymbol{A}$ in the prior by the position of the peak of the likelihood $\boldsymbol{A} \to \boldsymbol{h}$. Then we have along with (27), (21), and (18)

$$\begin{aligned} &\int \mathrm{dV}_{\boldsymbol{A}}\, p(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, m, I) p(\boldsymbol{A} \mid m, I) \\ &\approx p(\boldsymbol{A} = \boldsymbol{h} \mid m, I) \int \mathrm{dV}_{\boldsymbol{A}}\, p(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, m, I) \\ &\propto (2\pi)^{-\frac{m}{2}} \left(\frac{\hbar^2}{2}\right)^{-\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)\, e^{-\frac{\hat{\Phi}}{2\sigma^2}} \sigma^{-(N-m)}. \end{aligned}$$

Next we evaluate the integral over σ in (52), yielding the final result for the model probability

$$\begin{aligned} p\big(m \mid \boldsymbol{d}, I\big) &= \frac{\mathscr{Z}_m}{\mathscr{Z}}, \\ \mathscr{Z}_m &:= (2\pi)^{-\frac{m}{2}} \Gamma\left(\frac{m}{2}\right) \Gamma\left(\frac{N-m}{2}\right) \\ &\qquad \times \int \mathrm{dV}_{\omega}\; p(\boldsymbol{\omega} \mid I) \left(\frac{\hbar^2}{2}\right)^{-\frac{m}{2}} \left(\frac{\hat{\Phi}}{2}\right)^{-\frac{(N-m)}{2}}, \\ \mathscr{Z} &= \textstyle\sum_m \mathscr{Z}_m. \end{aligned}$$

A closer look reveals the competition between terms that tend to reduce the misfit and those terms (Ockham factor) that try to reduce the complexity of the model, i.e. keeping m small. This reveals one of the strengths of the Bayesian approach. In contrast to the so-called Akaike information criterion [1] which adds an ad-hoc penalty term to the likelihood, the Ockham factor is an integral part of the Bayesian approach.

The remaining integral is again evaluated by nested sampling [27]. Alternatively, MCMC-methods are also suited to solve such type of integrals [28].

In the previous sections of parameter estimation (Section 3.2) and signal estimation (Section 3.3), we have tacitly assumed that $m = 2(n_r + n_p)$ is known, which is not the case. Therefore, instead of (22) we actually have to compute the probability density for the parameters without knowing m, i.e. $p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid \boldsymbol{d}, I)$. This can be obtained in terms of (22) via the marginalization rule (1)

$$p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid \boldsymbol{d}, I) = \sum_m p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid \boldsymbol{d}, m, I)\, p(m \mid \boldsymbol{d}, I), \tag{53}$$

leading to a sum over m in (28). This means that we have to average the estimates also over m, e.g. $\sum_m \langle f_i \rangle\, p(m \mid \boldsymbol{d}, I)$ in (45). However, we will see in Sections 4.1 and 5.2.1 that $p(m \mid \boldsymbol{d}, I)$ is sharply peaked at a single value of m, i.e.

$$p(m \mid \boldsymbol{d}, I) \approx \delta_{m,\hat{m}},$$

where $\delta_{m,\hat{m}}$ is the Kronecker-delta. Then the sum in (53) shrinks to a single term and we have

$$p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid \boldsymbol{d}, I) \approx p(\boldsymbol{A}, \boldsymbol{\omega}, \sigma \mid \boldsymbol{d}, \hat{m}, I). \tag{54}$$

This means for the parameter estimates (Section 3.2) and signal estimation and uncertainty quantification (Section 3.3) that, once the most probable model $\hat{m}$ has been determined, it can be substituted for m throughout, as was done in the following sections.

4. Validation with artificial data

In this section, we want to test the efficiency of the approach with artificial data. These are generated with the help of the model function in (10) plus additive Gaussian noise with a level of $\sigma = 0.3$. The parameters of the model function are chosen at random and we use $m = 14$ or more specifically $n_r = 3$ and $n_p = 4$. A data set with $N = 5000$ data points was generated at equidistant times with a time step of $\frac{3}{4000}$ s. The generated noiseless data and their different sources are plotted in Fig. 1.

We apply the Bayesian approach to determine the probability for different models m and to estimate the parameters of the model function. Eventually, we decompose the signal into the pulsatile and respiratory part and provide uncertainties for all quantities.

4.1. Model selection

We start with model selection. As discussed in the introductory part, we restrict the discussion to the case $n_r \approx n_p$, or more precisely, $n_p = n_r$ for even $m/2$ and $n_p = n_r + 1$ for odd $m/2$. It should be remembered that m itself is always an even integer, since the basic functions cosine and sine are included in pairs. The results are displayed in Fig. 2, where we have varied $m = 8, 10, \ldots, 18$ corresponding to $(n_r, n_p) = (2, 2), (2, 3), \ldots, (4, 5)$.

As one can see, there is a sharp peak at $m = 14$ ($n_r = 3$, $n_p = 4$), which corresponds to the values that have been used to generate the artificial data. This result nicely illustrates the power of the Bayesian approach as far as model selection is concerned. A simple minimization of the misfit would always prefer the most complex model, i.e. the one with the largest number of higher harmonics. It is the Ockham factor that keeps m as small as possible.

4.2. Parameter estimation and signal analysis

Due to the sharp peak in the model probabilities, we can use (54) and simply continue with the most probable model. We then estimate the parameters of the most probable model function and decompose the signal into its source terms.

In Table 1, the estimated values for frequencies and the noise level are compared with the true values of the artificial data.

We observe that the parameters and uncertainties are reliably obtained by the Bayesian approach. We will not discuss the amplitudes, as it is more instructive to directly compare the signal and its individual contributions, which are depicted in Fig. 3. In the figure the true

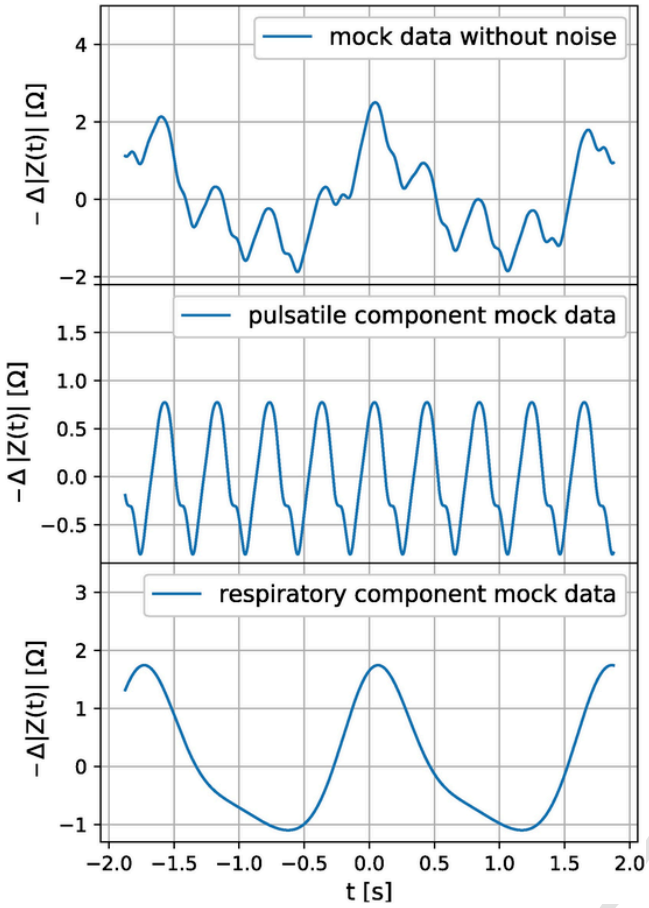

Fig. 1. Artificial data used to benchmark the approach. The upper panel shows the total signal. Pulsatile and respiratory part are plotted in the middle and lower panel, respectively. The ordinate axis stands for the negative absolute value of the impedance shifted by its mean value.

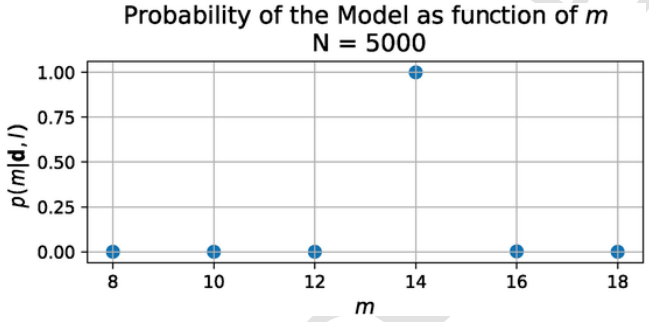

Fig. 2. Result of the model comparison for the artificial data. The probability for the model function is strongly peaked at $m = 14$ ($n_r = 3$, $n_p = 4$) which corresponds to the values that have been used to generate the artificial data.

curves (artificial data without noise) and the estimated signal and its components match within the line thickness.

4.3. Uncertainty quantification

In order to assess the reliability for our analysis, we determine the uncertainty bands of the signals. The results are depicted in Fig. 4.

The true total signal as well as its components lie within the uncertainty bands (see for comparison Fig. 1). The width of the uncertainty bands is very small due to the fact that we have way more data points than unknown parameters.

Table 1
Comparison of the estimated values and the true values of the frequencies and the noise level.

Estimates					
$\langle\omega_r\rangle$ [rad/s]	$\Delta\langle\omega_r\rangle$ [rad/s]	$\langle\omega_p\rangle$ [rad/s]	$\Delta\langle\omega_p\rangle$ [rad/s]	$\widehat{\sigma}_\omega$ [Ω]	$\Delta\widehat{\sigma}_\omega$ [Ω]
3.46	0.08	15.62	0.08	0.31	0.04
Artificial data input (truth)					
ω_r [rad/s]		ω_p [rad/s]		$\sigma[\Omega]$	
3.50		15.60		0.30	

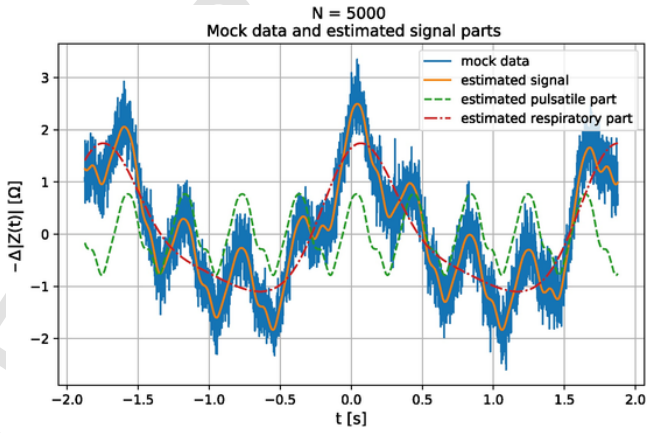

Fig. 3. Comparison of the artificial data signal with the Bayesian estimate and decomposition into pulsatile and respiratory part (including error bars). The estimated signal fits the data very well. In the figure the true curve (artificial data without noise) and the estimated signal cannot be distinguished. Meaning of ordinate axis as in Fig. 1.

4.4. Detecting non-stationary signals

Measurement artifacts or arrhythmic heartbeats will generally violate the implicit assumption of stationarity in the model function (9). While, at this level, our method is not yet suited to address strongly non-stationary signals produced by these anomalies, it is able to detect their occurrence. In order to test such a situation, we model a non-stationary signal by introducing a single arrhythmic heart beat with decreased amplitude. We show in Fig. 5 that such a 'local' non-stationarity can be exposed when plotting the misfit term (20) while moving the time window over the signal. A peak rises and falls above the base line exactly as the window slides over the arrhythmia, and can be clearly distinguished from the noise. The signal estimate cannot correctly reconstruct the arrhythmia due to the limitations of the model function and a lower limit for a reasonable time window size. This could be addressed with modifications to the model function, e.g. by equipping the harmonics with a phase, making the amplitudes time-dependent or augmenting the model function with Fermi functions. Such adaptions of the theory might be necessary to use it for other biomedical signals such as e.g. EMG signals.

5. Application to real bioimpedance data

Here, we apply the present Bayesian approach to analyze real EBI data that stems from a small number of volunteers. Ethical committee approval has been obtained. We use a set with $N = 5000$ data points and a time step of $\frac{3}{4000}$ s. We note that the method is agnostic to different signal compositions due to different experimental parameters such

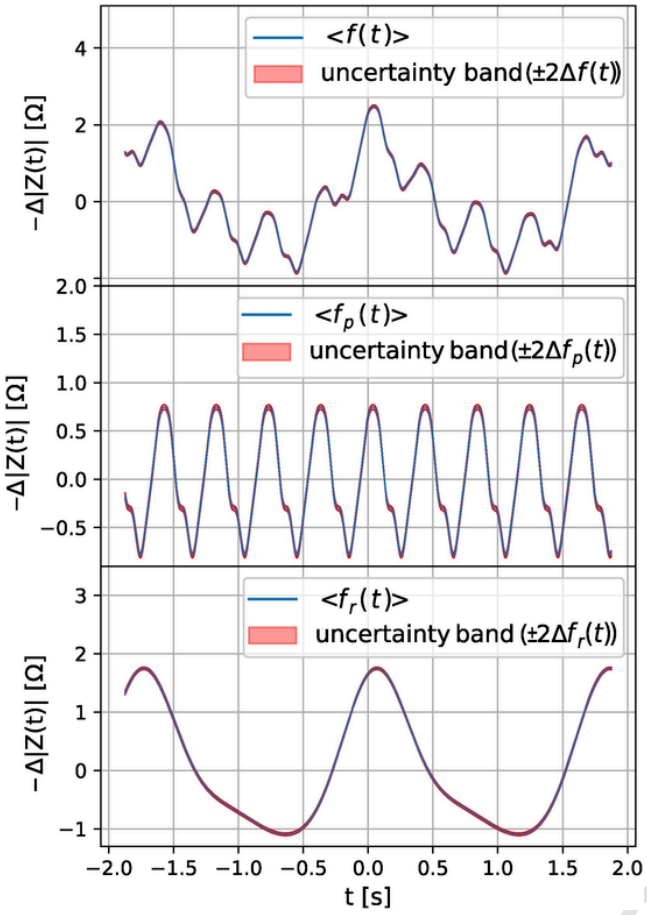

Fig. 4. Estimated signal and uncertainty band (uppermost panel) of the artificial data. The panel in the middle (bottom) shows the estimated pulsatile (respiratory) part and the corresponding uncertainty band. Meaning of ordinate axis as in Fig. 1.

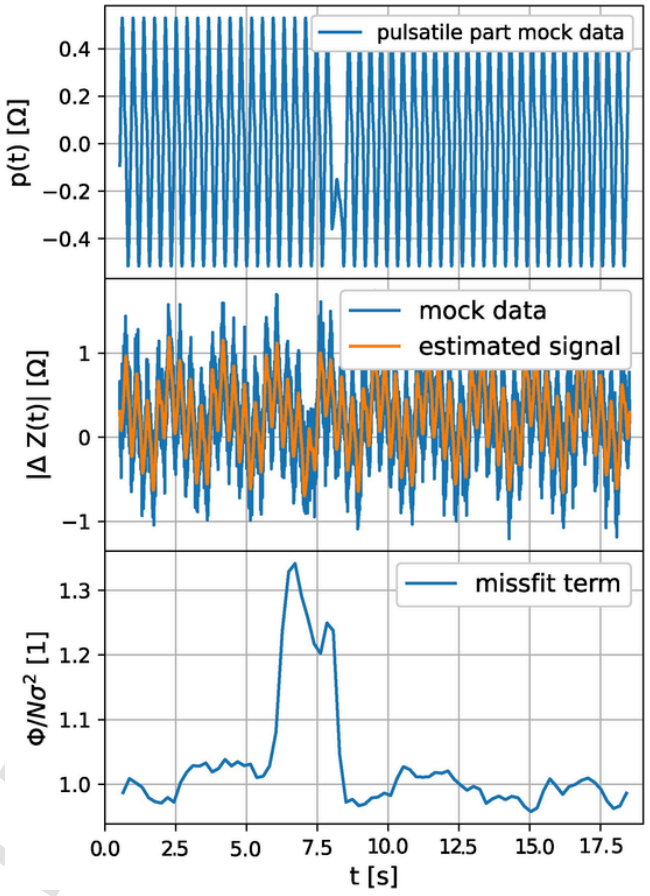

Fig. 5. Analysis of artificial data used to test a non-stationary signal. The upper panel shows the true pulsatile signal source with one pulse having decreased amplitudes to represent a heart arrhythmia. In the middle panel, this pulse is superimposed with a respiratory signal and noise (blue) and its reconstruction (orange, only intended for visual aid). The lower panel shows the normalized value of the misfit term (20) as a function of the moving time window midpoint (window width = 2 s).

as injection current amplitude and frequency. One must not apply any filters since that would violate above assumptions.

5.1. Elimination of the trend

For this measurement, one electrode was positioned on the left hand side of the hip and the other one on the right hand side of the neck. In the real data set, an additional long-term trend has been observed. Such long-term trends can arise due to, e.g., water displacements in the human body and can be modeled with additional terms. A convenient way to model the trend is to add a certain number of Legendre polynomials to the model function, as they are simple to implement and easy to normalize. This changes only the definition of the basis functions but has no impact on the other expressions we have derived. The new expansion reads

$$f(t) = \sum_{j=1}^{m} B_j \cdot G_j(t). \tag{55}$$

Here, the total number of basis functions m is now the sum of the number of the harmonic functions m_ω and the number of Legendre polynomials m_L. We use a uniform prior for m_L but with an upper limit $m_L^{\max}$ such that the Legendre polynomials are not able to describe the periodic signals we are interested in. This is easily achieved for the pulsatile part, which has the smallest period. For the respiratory part, the separation is not so obvious, and we will indeed see that it can lead to ambiguities in the case of data of poor quality.

5.2. High quality data

We begin with a data set with a high signal-to-noise ratio (SNR).

5.2.1. Model selection

First, we compute the probability for m_ω, the total number of higher harmonics, and for m_L, the number of polynomials required to cover the trend of the EBI signal. As before, since we have no preference as far as m_L concerned, we use a uniform prior. Then, the desired joint probability for the expansion orders $m = \{m_\omega, m_L\}$ is proportional to the marginal likelihood $p(\boldsymbol{d} \mid m, I)$, with the appropriate normalization. Results are shown in in Fig. 6.

Again we observe a pronounced maximum and we can, therefore, again use (54) with $\hat{m}$ standing for $n_r = 3$, $n_p = 4$, $m_L = 4$.

We also observe (not shown in the figure) that, when extending the time window of the data set, we need more Legendre polynomials to describe the trend properly.

5.2.2. Signal analysis

Next, we want to estimate the underlying signal of the EBI data and decomposition into its source terms. The result is displayed in Fig. 7.

We observe that the signal fits well into the error bars of the data. This is nice, but it only tells us that the model function is flexible enough to describe the data. A more interesting question concerns the assumption that the experimental data are distorted by additive uncor-

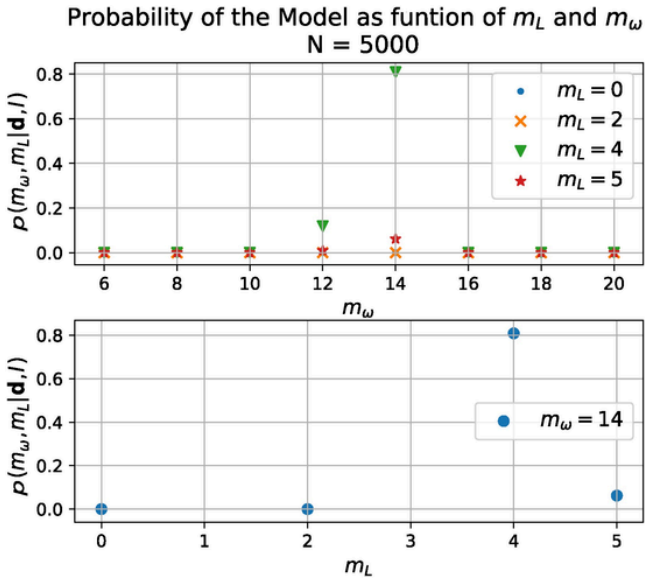

Fig. 6. Probability $p(m_\omega, m_L \mid \mathbf{d}, I)$ for the real data plotted for different values for the expansion orders m_ω and m_L. The upper part shows the dependence on m_ω for different values of m_L. The lower diagram represents the result as function of m_L for fixed $m_\omega = 14$. The model function with $m_\omega = 14$ and $m_L = 4$ is by far the most probable.

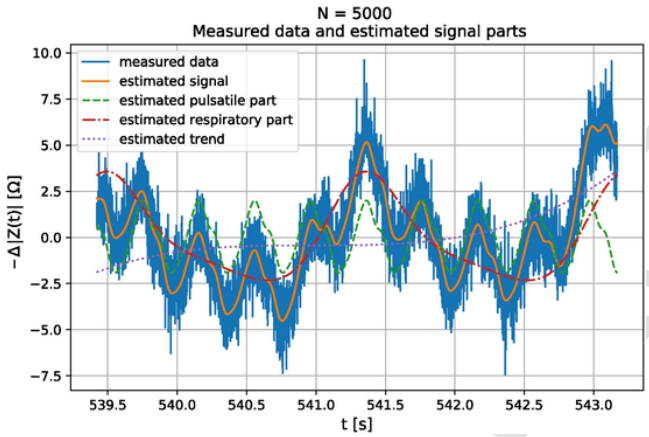

Fig. 7. Measured bioimpedance data, estimated signal and its components. Here, $N = 5000$, which corresponds approximately to a time interval of 4 s. Meaning of ordinate axis as in Fig. 1.

related Gaussian noise. To verify this assumption we take a closer look at the difference between the estimated signal and the measured data. We find that these differences are indeed normally distributed with a mean $\mu \approx 0$ and a standard deviation $\sigma \approx \widehat{\sigma}_\omega$. Hence, the assumption of Gaussian noise is corroborated by the data.

5.2.3. Uncertainty quantification

For the real data, it is particularly important to calculate the uncertainty bands. The estimated signal and its source terms along with the corresponding uncertainties can be seen in Fig. 8.

5.3. Noisy data

With the Bayesian approach we can also decompose signals with a very poor SNR. Such a noisy data set of a real EBI measurement signal is depicted in Fig. 9. It stems from a measurement where the electrodes were mounted on the lower right arm and on the right shoulder.

Also in this case we find a predominant probability for $n_r = 3$, $n_p = 4$, and $m_L = 4$. In Fig. 10 we compare the joint pdf for the fre-

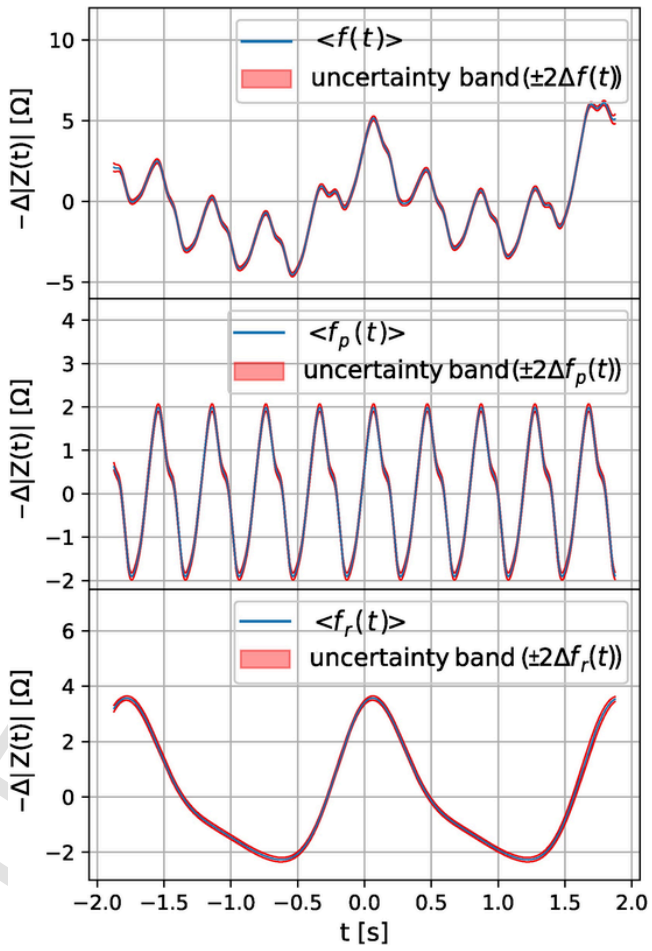

Fig. 8. The estimated signal and its uncertainty band for the real data in Fig. 7 is depicted in the upper panel. The middle (lower) panel shows the corresponding plots for the pulsatile (respiratory) part. Since the trend is not of primary interest, it is not shown here. Meaning of ordinate axis as in Fig. 1.

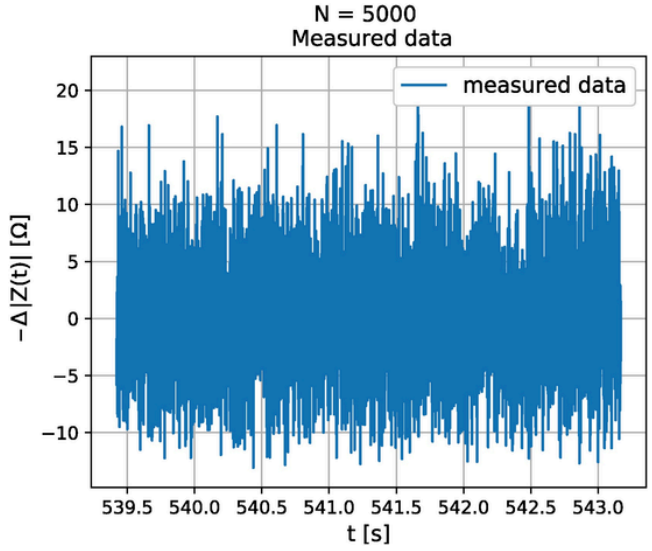

Fig. 9. A noisy data set from a real bioimpedance measurement. The number of data points N is again 5000, which corresponds approximately to an time interval of 4 s. Meaning of ordinate axis as in Fig. 1.

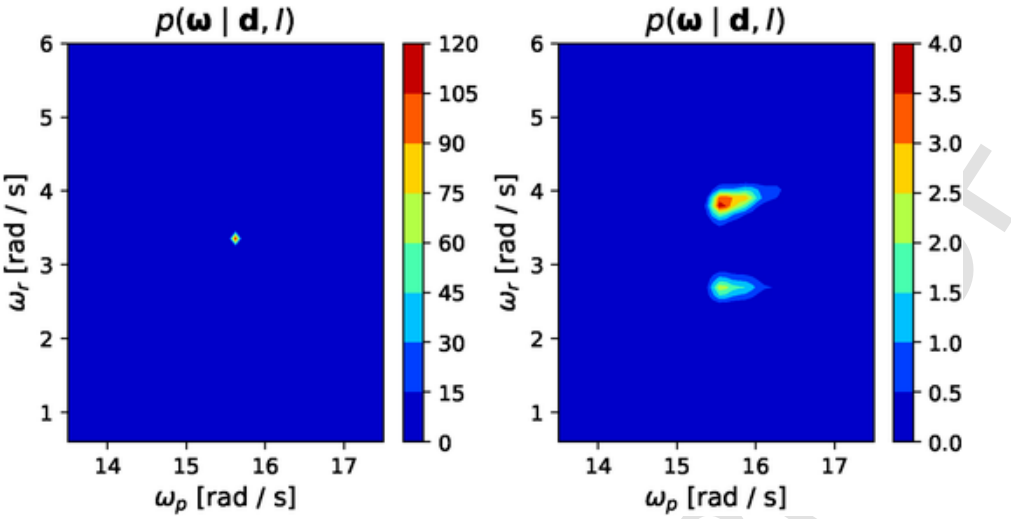

Fig. 10. Joint probability density for the frequencies for the data set of Fig. 7 (left plot) and those of the noisy data of Fig. 9 (right plot).

quencies $p(\omega_r, \omega_p \mid \boldsymbol{d}, I)$ for the data set with poor SNR with that of the previous data set, which had a significantly better SNR. In the latter case, we observe a single narrow peak, corresponding to well defined estimates for the frequencies with small uncertainty. In the low SNR data set, on the other hand, we find a bi-modal pdf. The two peaks are at the same position as far as ω_p is concerned but they correspond to two significantly different values for ω_r. Consequently, the marginal pdf for ω_p still allows a unique estimate of the pulsatile frequency with small uncertainty. The marginal pdf for ω_r, on the other hand, has two peaks, where the peak with the higher frequency at least has a significantly higher probability than the other one. A multi-model probability density for the respiratory frequency does not necessarily mean that the corresponding signal component cannot be reliably reconstructed. But in the present case the data are so noisy that the significance of the reconstructed respiratory signal is completely useless, as one can see in Fig. 11. It may surprise the reader that although the respiratory part of the spectrum is completely unreliable, the total signal still has a decent error band. The reason is the following: Respiratory part and trend cannot be distinguished any more, since – due to the considerable noise – no periodic long-wavelength signal can be detected. As a result, only the sum of the two signals can be reasonably well determined. The behavior of the pulsatile part is different, as it is restricted to higher frequencies by the prior, so that it cannot be misused to describe the slowly varying trend.

If one is indeed interested in the value of ω_r, only a much longer time window over several respiratory periods or better prior knowledge can help in this case. If one is, however, only interested in the pulsatile part of the signal, then even with the poor data available the signal can fairly convincingly be reconstructed, as can be observed in Fig. 11.

5.4. Detecting arrhythmias and artifacts

We apply our method to a particularly challenging real data set that contains a heart arrhythmia and an artifact right next to each other. Similar to Section 4.4, we recognize in Fig. 12 two subsequent peaks in the misfit term (bottom) that can be clearly distinguished form the noise, exactly where the artifact appears in the bioimpedance signal (middle) and where the arrhythmia is independently detected in invasive blood pressure measurements (A-line, top). More interestingly, the arrhythmia can be clearly distinguished from the artifact despite being less than one time window (2 s) apart. The arrhythmia is not correctly resolved in the signal reconstruction (see Section 4.4), however the artifact can be reconstructed by the Legendre polynomials.

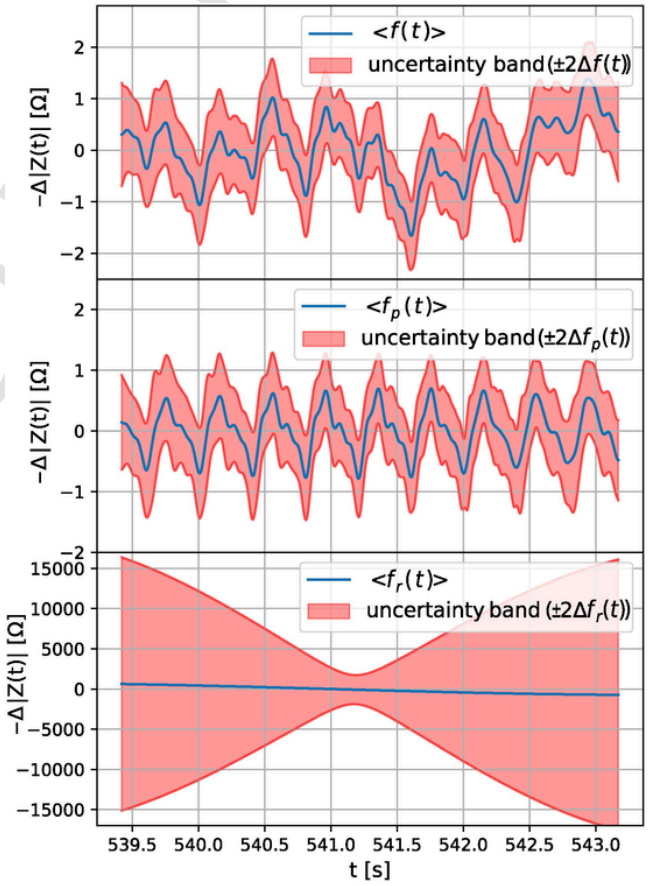

Fig. 11. Estimated signal and calculated uncertainty band (upper panel) for the noisy data of Fig. 9. The central panel shows the estimated pulsatile part and the corresponding uncertainty band. The bottom panel shows the estimated respiratory part and the corresponding uncertainty band. Since the long-term trend is not of primary interest it was not plotted here. Meaning of ordinate axis as in Fig. 1.

5.5. General findings

In the previous section, we have found that for both data sets the model with $n_r = 3$, $n_p = 4$, and $m_L = 4$ dominates the joint probability for these quantities. We have analyzed additional EBI data sets for dif-

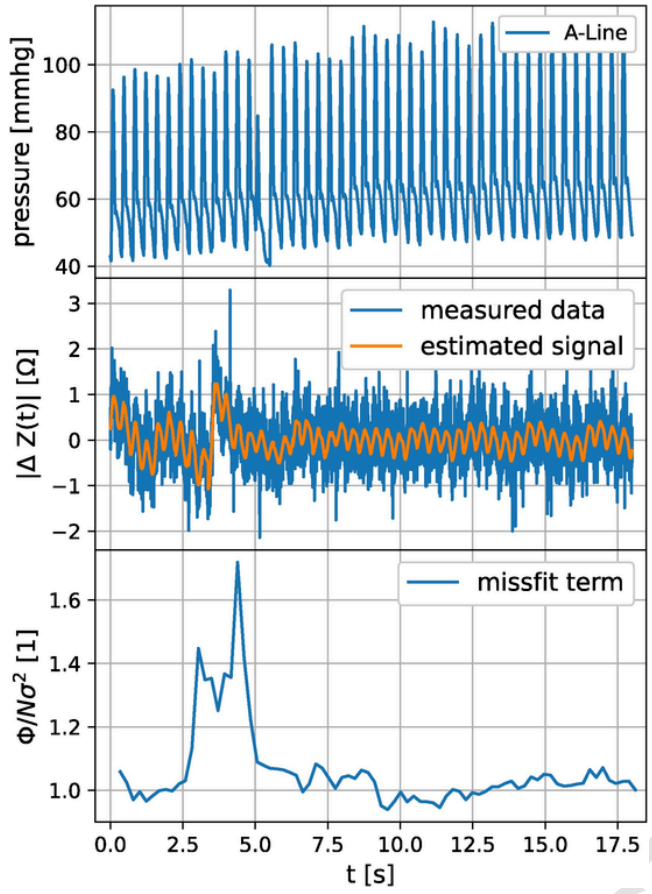

Fig. 12. Analysis of real data containing both an arrhythmia and an artifact, possibly produced by movement. The upper panel shows the invasively measured pressure (A-line) with an arrhythmic heartbeat. The middle panel shows the simultaneously measured bioimpedance signal (blue) with an artifact right before the arrhythmia and its reconstruction (orange, only intended for visual aid). The lower panel shows the normalized value of the misfit term (20) as a function of the moving time window midpoint.

ferent electrode positions. Most of these data sets lead to the same conclusion. Only in a few cases with poor SNR did the Bayesian analysis yield smaller values for n_p, which is understandable because data support for higher frequencies is lost in noise. Nevertheless, $m_\omega = 14$ can also be used in these cases, since the estimated amplitude of the additional functions turn out to be close to zero. Hence all data sets can be analyzed assuming a fixed set of expansion orders $n_r = 3$, $n_p = 4$ and $n_L = 4$.

In principle, it is not necessary that the probability for the expansion orders is peaked at a single point and that this point is the same for all datasets. The correct Bayesian approach is given in (53) as a sum over all values of the expansion orders. The advantage of one and the same single peak for all data sets consists 'merely' in the fact that the computational complexity is reduced significantly, as the sum over the expansion orders can be avoided (see (54)). Since the trend, which could come from movements of the patient or water displacements in the human body, is not predictable, it is impossible to say beforehand how many Legendre polynomials will be needed in the model function. When considering small time windows ($N \leq 5000$) only as in our case, 4 Legendre polynomials can sufficiently model the trend in the data.

The goal of this work is to decompose the signal into its components and therefore a rather small time window should be considered because of the natural variability in pulse frequency. On the other hand, it is clear that for a good estimation of the respiratory part, one should instead use larger time windows, where at least two respiration cycles are included. Time windows of about 4 s ($N \approx 5000$) appear to be a good compromise.

6. Discussion

We discuss conceptual differences and similarities to other methods of parameter estimation and signal reconstruction with EBI signals. The matter has been posed as an optimization problem, e.g. LASSO regression [5], principal component analysis (PCA) [22] or the estimation of a transfer function or filter function [19]. Formally this is

$$\{\boldsymbol{A}^{(\mathrm{opt})}, \boldsymbol{\omega}^{(\mathrm{opt})}\} = \arg\min_{\boldsymbol{A},\boldsymbol{\omega}} \mathscr{L}(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega})$$

with an optimality criterion or loss function $\mathscr{L}$. Amongst these, maximum likelihood (ML) estimates for the parameters are popular, particularly (regularized) least squares or those mentioned above. Since the noise is Gaussian and our prior is approximately flat, our parameter estimate for the amplitudes given the frequencies (36) is indeed equivalent to taking the argument of the minimized least squares for a given set of frequencies. However, in the Bayesian approach there is no need to introduce any unknown tuning parameters to regularize the, possibly ill-posed, optimization problem, and the method is robust by design. The analogy does not generally hold for the frequency estimation because the posterior probability density function for the frequencies (28) is not Gaussian and can be skewed or even multi-modal (Fig. 10). Most importantly, the Bayesian approach allows to prescribe a priori known frequency ranges and then demands to integrate over the probability density function instead of optimizing it. I.e., there appears no optimization problem at all, neither unconstrained nor constrained by any frequency ranges. A drawback can be that the resulting integrals can be difficult to solve, too. Another interesting difference is found in the signal reconstruction or decomposition. Inserting the parameter estimate into the model function,

$$f(t \mid \mathbf{A}^{(\mathrm{opt})}, \omega^{(\mathrm{opt})}),$$

is not the same as the Bayesian result (45). The integral takes into account all available information, e.g. skewness, asymmetry or multi-modality of the likelihood (acting as a 'loss function' in the above sense). This also allows to quantify the uncertainties of the signal (source) estimate (47) on the same probabilistic footing too, in contrast to point estimates resulting from optimization. Finally, we do not need ad-hoc procedures or diagnostics to prevent over-fitting in the model selection. This is inherently taken care of by the method, as discussed in Section 3.4.

7. Conclusion

In this paper, the decomposition of the measured EBI signals into their components, pulsatile and respiratory part, was derived within the framework of Bayesian probability theory. This allows a consistent treatment of the experimental uncertainties and to include physiological or medical prior knowledge about the various quantities such as the natural frequency range of the periodic signal parts. Moreover and on the same footings, the probabilistic approach allows to separate the signal of interest from disturbing influences from other sources that may lead to long term trends, and to infer the uncertainties of the recovered signal and its components. The formalism was tested with artificial data and also with data from real EBI measurements. It was shown, that it allows for an efficient decomposition of the signal into its components without filters and without data or information from further measurements. Due to the rather large number of data points N within the considered time window as compared to the number of unknown model parameters, the uncertainty bands of the estimated signal and of all the components are usually quite narrow. Thus, the estimated signals and components are reliable for the typical EBI data sets. Even in the case of very poor data, where nothing is recognizable to the naked eye, the

signal and its pulsatile component can be recovered with a reasonable uncertainty. The time window was chosen such that only a few heart beats are included to allow for a fast reconstruction and to avoid unnecessary uncertainties due to changes in the frequency of pulsatile and respiratory part. By shifting this time window over the time series of the measured data, one can monitor changes of these frequencies. Since the EBI signals used were measured on sedated persons, motion artifacts were not considered in this work.

The ongoing research pursues two aspects, robustness against all kinds of artifacts and the combined treatment of data from multiple electrode pairs. The method can be adapted for other biomedical signals such as electrocardiogram (ECG) or electroencephalography (EEG).

Author contributions

Christof Pichler: Investigation, software, methodology, validation, formal analysis, visualization, writing – original draft. Sascha Ranftl: Investigation, methodology, conceptualization, formal analysis, writing – original draft, writing – review & editing, project administration. Arnulf Heller: Data curation, resources. Enrico Arrigoni: Conceptualization, supervision, writing – review & editing. Wolfgang von der Linden: Formal analysis, conceptualization, funding acquisition, supervision, writing – review & editing.

Conflict of interest

"ImPress MedTech GmbH" owns the real experimental data and shared the data for analysis, respectively for applying our method to. This data was collected by co-author Arnulf Heller with Bernd Höfler, who are (co-)founder, (co-)manager, (co-)shareholder of the company and industry partner "ImPress MedTech GmbH". The other authors, Christof Pichler, Sascha Ranftl, Enrico Arrigoni and Wolfgang von der Linden are with the Institute of Theoretical and Computational Physics of Graz University of Technology, and did not receive any funding external of Graz University of Technology. These 4 authors particularly did not receive compensation from ImPress MedTech GmbH, they were financially supported from Graz University of Technology only. The authors had full disclosure of all necessary measurement details. The funding sources were not involved in the development of the method/theory, nor in the study design, nor in the analysis or interpretation of the data other than sharing the data; nor in the writing of the report; nor in the decision to submit the article for publication.

Funding

This work was in part financially supported by the Lead Project 'Mechanics, Modeling and Simulation of Aortic Dissection' of Graz University of Technology and ImPress MedTech GmbH.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

The authors would like to thank ImPress MedTech GmbH for providing the measurement data and Bernd Höfler for recording patient data in Intensive Care Unit at Medical University Hospital Graz. We thank Vahid Badeli, Alice Reinbacher-Köstinger and the anonymous reviewers for their comments. We acknowledge support by TUG Open Access Publishing Fund.

Appendix A. Proof of Eqs. (36) and (37)

Here we prove Eqs. (36) and (37). We begin with the expression for the mean of the amplitudes. For simplicity, we suppress the argument m in the conditional complex

$$\langle A_k \rangle_{\boldsymbol{\omega}} = \int \mathrm{dV}_{\boldsymbol{A}}\, A_k\, p\,(\boldsymbol{A} \mid \boldsymbol{\omega}, \boldsymbol{d}, I)\,. \tag{56}$$

Next we use (35) and obtain

$$\langle A_k \rangle_{\boldsymbol{\omega}} = \int d\sigma\, \langle A_k \rangle_{\boldsymbol{\omega},\sigma}\, p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, I) \tag{57}$$

with

$$\langle A_k \rangle_{\boldsymbol{\omega},\sigma} = \int \mathrm{dV}_{\boldsymbol{A}}\, A_k\, p\,(\boldsymbol{A} \mid \boldsymbol{d}, \boldsymbol{\omega}, \sigma, I)\,. \tag{58}$$

For the last pdf we use Bayes' theorem

$$p\,(\boldsymbol{A} \mid \boldsymbol{d}, \boldsymbol{\omega}, \sigma, I) = \frac{1}{\boldsymbol{\mathcal{Z}}} p\,(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, I)\; p\,(\boldsymbol{A} \mid \boldsymbol{\omega}, \sigma, I)\,.$$

As argued before in (52) or (28),we can replace $\boldsymbol{A}$ in the prior with $\widehat{\boldsymbol{A}} = \boldsymbol{h}$, i.e. with the peak position in the likelihood. That makes the prior a constant with respect to $\boldsymbol{A}$, so we actually have

$$p(\boldsymbol{A} \mid \boldsymbol{d}, \boldsymbol{\omega}, \sigma, I) = \frac{1}{\boldsymbol{\mathcal{Z}}'} p(\boldsymbol{d} \mid \boldsymbol{A}, \boldsymbol{\omega}, \sigma, I).$$

From (21), (13), and (17) we obtain that the likelihood in terms of $\boldsymbol{A}$ is a Gaussian, centered at $\widehat{\boldsymbol{A}} = \boldsymbol{h} = H^T \boldsymbol{d}$ and with a diagonal covariance

$$\langle \Delta A_k \Delta A_{k'} \rangle_{\boldsymbol{\omega},\sigma} = \delta_{\mathrm{kk}'} \sigma^2, \tag{59}$$

and $\langle A_{\mathrm{k}} \rangle_{\omega,\sigma} = h_{\mathrm{k}}$. Along with the proper normalization, we have

$$p(\boldsymbol{A} \mid \boldsymbol{d}, \boldsymbol{\omega}, \sigma, I) = (2\pi)^{-\frac{m}{2}} \sigma^{-m} \exp\left(-\frac{1}{2\sigma^2}(\boldsymbol{A} - \boldsymbol{h})^2\right).$$

Then (57) leads to

$$\langle A_k \rangle_{\boldsymbol{\omega}} = \int d\sigma\, \langle A_k \rangle_{\boldsymbol{\omega},\sigma}\, p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, I) = h_k. \tag{60}$$

Similarly, we obtain

$$\langle \Delta A_k \Delta A_{k'} \rangle_{\boldsymbol{\omega}} = \int d\sigma\, \langle \Delta A_k \Delta A_{k'} \rangle_{\boldsymbol{\omega},\sigma}\, p(\sigma \mid \boldsymbol{\omega}, \boldsymbol{d}, I). \tag{61}$$

With (59) for the conditional covariance, we finally find the desired result

$$\langle \Delta A_k \Delta A_{k'} \rangle_{\boldsymbol{\omega}} = \delta_{\mathrm{kk}'} \left\langle \sigma^2 \right\rangle_{\boldsymbol{\omega}}. \tag{62}$$

Along with (33) we can replace the last expression by

$$\langle \Delta A_k \Delta A_{k'} \rangle_{\boldsymbol{\omega}} = \delta_{\mathrm{kk}'} \widehat{\sigma}^2_{\boldsymbol{\omega}}. \tag{63}$$

Appendix B. Second moment of the signal

Here we outline the evaluation of the second moment of the signal. According to (41), i.e.

$$\left\langle f_i^2 \right\rangle = \int \mathrm{dV}_{\boldsymbol{\omega}}\, p\,(\boldsymbol{\omega} \mid \boldsymbol{d}, I)\;\; \left\langle (H\boldsymbol{A})_i^2 \right\rangle_{\boldsymbol{\omega}}, \tag{64}$$

we need the conditional second moment $\left\langle (H\boldsymbol{A})_i^2 \right\rangle_{\boldsymbol{\omega}}$. Due to the Gaussian form of $p(\boldsymbol{A} \mid \widehat{\sigma}_{\omega}, \boldsymbol{\omega}, \boldsymbol{d}, I)$ specified in (43) with mean $\langle \boldsymbol{A} \rangle_{\boldsymbol{\omega}} = \boldsymbol{h}$ and covariance $\langle \Delta A_l \Delta A_{l'} \rangle_{\boldsymbol{\omega}} = \delta_{\mathrm{ll}'} \widehat{\sigma}^2_{\boldsymbol{\omega}}$ we obtain

$$\begin{aligned} \left\langle \left(H\boldsymbol{A}\right)_i^2 \right\rangle_{\boldsymbol{\omega}} &= \textstyle\sum_{\mathrm{ll}'} H_{\mathrm{il}} H_{\mathrm{il}'} \left\langle A_l A_{l'} \right\rangle_{\boldsymbol{\omega}} \\ &= \textstyle\sum_{\mathrm{ll}'} H_{\mathrm{il}} H_{\mathrm{il}'} \left(\langle A_l \rangle_{\boldsymbol{\omega}} \langle A_{l'} \rangle_{\boldsymbol{\omega}} + \langle \Delta A_l \Delta A_{l'} \rangle_{\boldsymbol{\omega}} \right) \\ &= \left(H\boldsymbol{h}\right)_i^2 + \widehat{\sigma}^2_{\boldsymbol{\omega}} \left(HH^T\right)_{\mathrm{ii}}. \end{aligned} \tag{65}$$

Insertion into (64) yields

$$\left\langle f_i^2 \right\rangle = \int \mathrm{dV}_{\boldsymbol{\omega}} \left[(H\boldsymbol{h})_i^2 + \hat{\sigma}_\omega^2 \left(\mathrm{HH}^T\right)_{\mathrm{ii}} \right] p(\boldsymbol{\omega} \mid \boldsymbol{d}, I). \tag{66}$$

The squared uncertainty, denoted by $(\Delta f_i)^2$, is defined as

$$(\Delta f_i)^2 = \left\langle f_i^2 \right\rangle - \langle f_i \rangle^2 . \tag{67}$$

Then along with (44) and (45) we obtain

$$\begin{aligned} (\Delta f_i)^2 \quad = \quad & \int \mathrm{dV}_{\boldsymbol{\omega}}\, p\left(\boldsymbol{\omega} \mid \boldsymbol{d}, I\right) \\ & \times \left[\left\langle \left[\left(H\boldsymbol{h}\right)_i - \left\langle \left(H\boldsymbol{h}\right)_i \right\rangle_{\boldsymbol{\omega}} \right]^2 \right\rangle_{\boldsymbol{\omega}} + \left(\mathrm{HH}^T\right)_{\mathrm{ii}} \hat{\sigma}_{\boldsymbol{\omega}}^2 \right]. \end{aligned} \tag{68}$$

Appendix C. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.bspc.2021.102541.

References

[1] H. Akaike, Prediction and entropy, in: A.C. Atkinson, S.E. Fienberg (Eds.), A Celebration of Statistics, Springer, New York, 1985, pp. 1–24, doi:10.1007/978-1-4613-8560-8_1.

[2] J.M. Bernardo, Reference posterior distributions for Bayesian inference, J. R. Stat. Soc. B (Stat. Methodol.) 41 (1979) 113–147. www.jstor.org/stable/2985028.

[3] L.G. Bretthorst, Bayesian Spectrum Analysis and Parameter Estimation, Springer Science & Business Media, 2013, doi:10.1007/978-1-4684-9399-3.

[4] B.H. Brown, D.C. Barber, A. Morice, A.D. Leathard, Cardiac and respiratory related electrical impedance changes in the human thorax, IEEE Trans. Biomed. Eng. 41 (1994) 729–734, doi:10.1109/10.310088.

[5] M. Butsenko, O. Martens, A. Krivoshei, Y. Le Moullec, Sparse reconstruction method for separating cardiac and respiratory components from electrical bioimpedance measurements, Elektron. Elektrotech. 24 (2018) 57–61, doi:10.5755/j01.eie.24.5.21844.

[6] D.H. Diaz, Ó. Casas, R. Pallas-Areny, Heart rate detection from single-foot plantar bioimpedance measurements in a weighing scale, 2010 Conf. Proc. IEEE Eng. Med. Biol. Soc., 2010, pp. 6489–6492, doi:10.1109/IEMB.S.2010.5627358.

[7] C. El-Hajj, P. Kyriacou, A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure, Biomed. Signal Process. Control 58 (2020) 101870, doi:10.1016/j.bspc.2020.101870.

[8] S. Grimnes, O.G. Martinsen, Bioimpedance. Wiley Encyclopedia of Biomedical Engineering, 2006, doi:10.1002/9780471740360.ebs0128.

[9] W. Hu, H. Sun, X. Wang, A study on methods for impedance cardiography, 1997 Conf. Proc. IEEE Eng. Med. Biol. Soc., 1997, pp. 2074–2077, doi:10.1109/IEMBS.1997.758758.

[10] D.G. Jakovljevic, M.I. Trenell, G.A. MacGowan, Bioimpedance and bioreactance methods for monitoring cardiac output, Best. Pract. Res. Clin. Anaesthesiol. 28 (2014) 381–394, doi:10.1016/j.bpa.2014.09.003.

[11] E.T. Jaynes, Information theory and statistical mechanics, Phys. Rev. 106 (1957) 620–630, doi:10.1103/PhysRev.106.620.

[12] E.T. Jaynes, Information theory and statistical mechanics. II, Phys. Rev. 108 (1957) 171–190, doi:10.1103/PhysRev.108.171.

[13] E.T. Jaynes, Prior probabilities, IEEE Trans. Syst. Sci. Cybern. 4 (1968) 227–241, doi:10.1109/TSSC.1968.300117.

[14] E.T. Jaynes, Probability Theory: The Logic of Science, Cambridge University Press, 2003, doi:10.1017/CBO9780511790423.

[15] J. Kapur, H. Kesavan, Entropy optimization principles and their applications, in: V.P. Singh, M. Fiorentino (Eds.), Entropy and Energy Dissipation in Water Resources, Springer Netherlands, 1992, pp. 3–20, doi:10.1007/978-94-011-2430-0_1.

[16] M.G. Kendall, P.A. Moran, Geometrical Probability, 1963. https://archive.org/details/geometricalproba033077mbp/page/n11/mode/2up.

[17] A. Krivoshei, A bio-impedance signal synthesiser (BISS) for testing of an adaptive filtering system, 2006 International Biennial Baltic Electronics Conference, 2006, pp. 1–4, doi:10.1109/BEC.2006.311104.

[18] A. Krivoshei, M. Min, T. Parve, A. Ronk, An adaptive filtering system for separation of cardiac and respiratory components of bioimpedance signal, 2006 IEEE Int. Workshop Med. Meas. App., 2006, pp. 10–15, doi:10.1109/MEMEA.2006.1644448.

[19] A. Krivoshei, H. Uuetoa, M. Min, P. Annus, T. Uuetoa, J. Lamp, CAP waveform estimation from the measured electrical bioimpedance values: patient's heart rate variability analysis, 2015 Conf. Proc. IEEE Eng. Med. Biol. Soc., 2015, pp. 2788–2791, doi:10.1109/EMBC.2015.7318970.

[20] W. von der Linden, V. Dose, U. von Toussaint, Bayesian Probability Theory: Applications in the Physical Sciences, Cambridge University Press, 2014.

[21] M. Min, T. Parve, A. Kink, Thoracic bioimpedance as a basis for pacing control, Ann. N. Y. Acad. Sci. 873 (1999) 155–166, doi:10.1111/j.1749-6632.1999.tb09463.x.

[22] Y.M. Mughal, A. Krivoshei, P. Annus, Separation of Cardiac and Respiratory Components from the Electrical Bio-Impedance Signal Using PCA and Fast ICA, 2013 arXiv:1307.0915.

[23] L. Peter, N. Noury, M. Cerny, A review of methods for non-invasive and continuous blood pressure monitoring: pulse transit time method is promising?, IRBM 35 (2014) 271–282, doi:10.1016/j.irbm.2014.07.002.

[24] R. Pikkemaat, S. Lundin, O. Stenqvist, R.D. Hilgers, S. Leonhardt, Recent advances in and limitations of cardiac output monitoring by means of electrical impedance tomography, Anesth. Analg. 119 (2014) 76–83, doi:10.1213/ANE.0000000000000241.

[25] J. Rosell, K.P. Cohen, J.G. Webster, Reduction of motion artifacts using a two-frequency impedance plethysmograph and adaptive filtering, IEEE Trans. Biomed. Eng. 42 (1995) 1044–1048, doi:10.1109/10.464380.

[26] D. Sivia, J. Skilling, Data Analysis: A Bayesian Tutorial, Oxford University Press, 2006.

[27] J. Skilling, Nested sampling for general Bayesian computation, Bayesian Anal. 1 (2006) 833–860, doi:10.1214/06-BA127.

[28] U. von Toussaint, Bayesian inference in physics, Rev. Mod. Phys. 83 (2011) 943–999, doi:10.1103/RevModPhys.83.943.