



**HAL**  
open science

# Precise periodic components estimation for chronobiological signals through Bayesian Inference with sparsity enforcing prior

Mircea Dumitru, Ali Mohammad-Djafari, Simona Baghai Sain

## ► To cite this version:

Mircea Dumitru, Ali Mohammad-Djafari, Simona Baghai Sain. Precise periodic components estimation for chronobiological signals through Bayesian Inference with sparsity enforcing prior. EURASIP Journal on Bioinformatics and Systems Biology, 2016, 1, pp.548 - 548. 10.1186/s13637-015-0033-6 . hal-01475255

**HAL Id: hal-01475255**

**<https://hal.science/hal-01475255v1>**

Submitted on 23 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RESEARCH

Open Access



# Precise periodic components estimation for chronobiological signals through Bayesian Inference with sparsity enforcing prior

Mircea Dumitru<sup>1,2\*</sup>, Ali Mohammad-Djafari<sup>1</sup> and Simona Baghai Sain<sup>1,3</sup>

## Abstract

The toxicity and efficacy of more than 30 anticancer agents present very high variations, depending on the dosing time. Therefore, the biologists studying the circadian rhythm require a very precise method for estimating the periodic component (PC) vector of chronobiological signals. Moreover, in recent developments, not only the dominant period or the PC vector present a crucial interest but also their stability or variability. In cancer treatment experiments, the recorded signals corresponding to different phases of treatment are short, from 7 days for the synchronization segment to 2 or 3 days for the after-treatment segment. When studying the stability of the dominant period, we have to consider very short length signals relative to the prior knowledge of the dominant period, placed in the circadian domain. The classical approaches, based on Fourier transform (FT) methods are inefficient (i.e., lack of precision) considering the particularities of the data (i.e., the short length). Another particularity of the signals considered in such experiments is the level of noise: such signals are very noisy and establishing the periodic components that are associated with the biological phenomena and distinguishing them from the ones associated with the noise are difficult tasks. In this paper, we propose a new method for the estimation of the PC vector of biomedical signals, using the biological prior informations and considering a model that accounts for the noise. The experiments developed in cancer treatment context are recording signals expressing a limited number of periods. This is a prior information that can be translated as the sparsity of the PC vector. The proposed method considers the PC vector estimation as an Inverse Problem (IP) using the general Bayesian inference in order to infer the unknown of our model, i.e. the PC vector but also the hyperparameters (i.e the variances). The sparsity prior information is modeled using a sparsity enforcing prior law. In this paper, we propose a Student's  $t$  distribution, viewed as the marginal distribution of a bivariate normal-inverse gamma distribution. We build a general infinite Gaussian scale mixture (IGSM) hierarchical model where we assign prior distributions also for the hyperparameters. The expression of the joint posterior law of the unknown PC vector and hyperparameters is obtained via Bayes rule, and then, the unknowns are estimated via joint maximum a posteriori (JMAP) or posterior mean (PM). For the PM estimator, the expression of the posterior distribution is approximated by a separable one, via variational Bayesian approximation (VBA), using the Kullback-Leibler (KL) divergence. For the PM estimation, two possibilities are considered: an approximation with a partially separable distribution and an approximation with a fully separable one. Both resulting algorithms corresponding to the PM estimation and the one corresponding to the JMAP estimation are iterative algorithms. The algorithms are presented in detail and are compared with the ones corresponding to the Gaussian model. We examine the

(Continued on next page)

\*Correspondence: Mircea.Dumitru@lss.supelec.fr

<sup>1</sup>Laboratoire des signaux et systèmes (L2S), UMR 8506 CNRS–CentraleSupélec–Univ. Paris-Sud, CentraleSupélec, Plateau de Moulon, 91192 Gif-sur-Yvette, France

<sup>2</sup>Rythmes Biologiques et Cancers (RBC), UMR 776 INSERM–Univ. Paris-Sud, Campus CNRS, 94801 Villejuif, France

Full list of author information is available at the end of the article

(Continued from previous page)

convergency of the algorithms and give simulation results to compare their performances. Finally, we show simulation results on synthetic and real data in cancer treatment applications. The real data considered in this paper examines the rest-activity patterns of KI/KI Per2::luc mouse, aged 10 weeks, singly housed in RealTime Biolumicorder (RT-BIO).

**Keywords:** Periodic component (PC) vector estimation, Sparsity enforcing, Bayesian parameter estimation, Variational Bayesian approximation (VBA), Kullback-Leibler (KL) divergence, Infinite Gaussian scale mixture (IGSM), Normal-inverse gamma, Inverse problem, Joint maximum a posteriori (JMAP), Posterior mean (PM), Chronobiology, Circadian rhythm, Cancer treatment

## 1 Introduction

Several biological processes in living organisms follow oscillations that repeat themselves about every 24 h—these oscillations are called circadian rhythms and together with other periodic phenomena, they are the object of study of chronobiology [1–3]. In mammals, circadian rhythms involve all organs, tissues, and cells and are supervised by the circadian timing system (CTS), a set of molecular clock genes that cross-regulate each other by positive and negative feedback loops [4–6]. More precisely, the CTS consists of a central pacemaker, the suprachiasmatic nuclei (SCN) in the hypothalamus, which is made sensitive to light by retinal afferents and which coordinates the molecular clocks in the peripheral organs by releasing diffusible and neurophysiological signals [3]. The period of the CTS, which is about 24 h, is therefore regularly calibrated by the succession of light and day and can be influenced by other environmental factors, such as socio-professional interactions and feeding times [5]. The resulting circadian physiologic fluctuations are observed in sleep-wakefulness and rest-activity alternations, body temperature, cortisol secretion by the adrenal gland, and melatonin secretion by the pineal gland, and they involve as well the sympathetic and the parasympathetic systems [6].

Former studies have already shown how taking chronobiology into account can improve anticancer treatment efficacy and reduce at the same time their toxicity (increasing therefore their tolerability), contrary to the previous “the worst the toxicity, the better the efficacy” paradigm [7–10]. The molecular clocks are involved in the regulation of important processes such as cell cycle and proliferation, DNA damage sensing and repair, apoptosis, angiogenesis, pharmacodynamics, and pharmacokinetics; therefore, they can greatly influence the metabolism, transportation, and detoxification of drugs [11].

Tolerability to anticancer treatments has been proven to depend significantly on their timing in respect to the circadian rhythms, measuring up to tenfold changes in the tolerability to drug administration at different circadian times for 40 anticancer drugs in rodents and up to fivefold in patients [11, 12]. Notably, chemotherapeutic agents proved to be at their best efficacy, both administered alone and combined, when they are also at

their best tolerability level, i.e., when they are least toxic to the healthy tissues. Furthermore, relevant interpatient variability in circadian rhythms have been observed and can be due to factors such as gender, age, and genetic polymorphisms; therefore, anticancer drugs dosing and timing need to be personalized, at least for subtypes of patients with similar chronotoxicity key features. Modulating drugs administration according to the patient's circadian rhythms is known as chronotherapy [13, 14]. On the other hand, administrating anticancer drugs at their most toxic time causes the disruption of molecular clocks synchronization, which has been shown to accelerate the cancer evolution [15–20].

In order to optimize cancer treatment, once proven that a certain drug effects are susceptible to circadian rhythms, we want to identify its best administration time. First, for each drug is proved the correlation with the circadian rhythms in a rodent model, which has been proved to well represent the human circadian physiology [11]. This is achieved by studying the chronotoxicity of the drug, inferred by body weight loss and histopathologic lesions, at different circadian times (CT or ZT, from Zeitgeber time). The mice circadian clock is synchronized by exposure to light for 12 h, followed by 12 h of dark, repeating this cycle and its rhythm is detected by tracking the expression of one or more of its core genes (normally Bmal1, Per2, Rev-erb $\alpha$ , or Clock are used). Mice with a disrupted clock (clock-defective mice, obtained via the functional knock out of one of its genes, normally Per2) are used to confirm the relevance of the molecular clock for the drug toxicity. At the same time, the main characteristics of the circadian expression of these observed genes are studied to observe whether the administration of the drug modifies them.

Once defined the CTs at which the drug best and worst tolerability is observed, we can look for the molecular mechanisms that influence it. Genes influencing the pharmacokinetics (absorption, distribution, metabolism, and excretion) of the drug are a good starting point, and we can follow how their expression correlates with the higher or lower drug chronotoxicity. For instance, the transporter abcc2, involved in the cellular efflux of several drugs, has been shown to influence irinotecan

chronotolerance in ileum, according to the circadian changes in *abcc2* local expression [21]. The circadian clocks of the mice used in the experiments whose data we analyse are first synchronized to the same day-night alternation where 12 h of light are followed by 12 h of dark (LD12:12). After synchronization, the mice are kept in constant darkness (DD), which implies the subtraction of the light. Throughout the experience, gene expression and rest-activity are measured to establish how the basic parameters of their circadian rhythms (period, acrophase, amplitude) vary in respect to the drug treatment. Both measures are allowed by an innovative monitoring device, the RealTime-Biolumicorder (RT-BIO) [22]. The locomotor activity is detected by an infrared sensor, whereas the gene expression is measured at the post-translational level in mice engineered to express the gene of interest together with luciferase (fLUC), so that the gene activity is marked by bioluminescence detected by a photomultiplier tube. Common mouse strains used are C57BL/6-based [7, 21] and 129S1/SvImJ [23]. The acrophase and amplitude depend on the periodic component (PC) vector, so a major interest is the study of the periodicity of such time series, i.e., the estimation of the PC vector and the stability or the variability of the dominant period, requiring a precise PC vector variation analysis. The periodical phenomena were studied with different approaches in different particular conditions [24–40] using in general fast Fourier transform (FFT)-based methods. The major limitation when studying such data is given by their reduced length, due to the duration of the experiments. The objective of an accurate description of the periodic components variation during the experiments can be formulated as the need of a method that can give a precise estimation of the PC vector from a limited number of data. Also, the method must be able to distinguish the peaks from the PC vector due to the biological phenomena and the peaks due to the measurements errors. The real data considered in this article is a chronobiological time series, measuring the locomotor activity. In order to observe the variation or the stability of the dominant periods, very short intervals of the recorded time series are considered. The prior knowledge is the presence of the circadian rhythm: the PC vector is sparse, having a limited number of non-zero elements, inside the circadian interval.

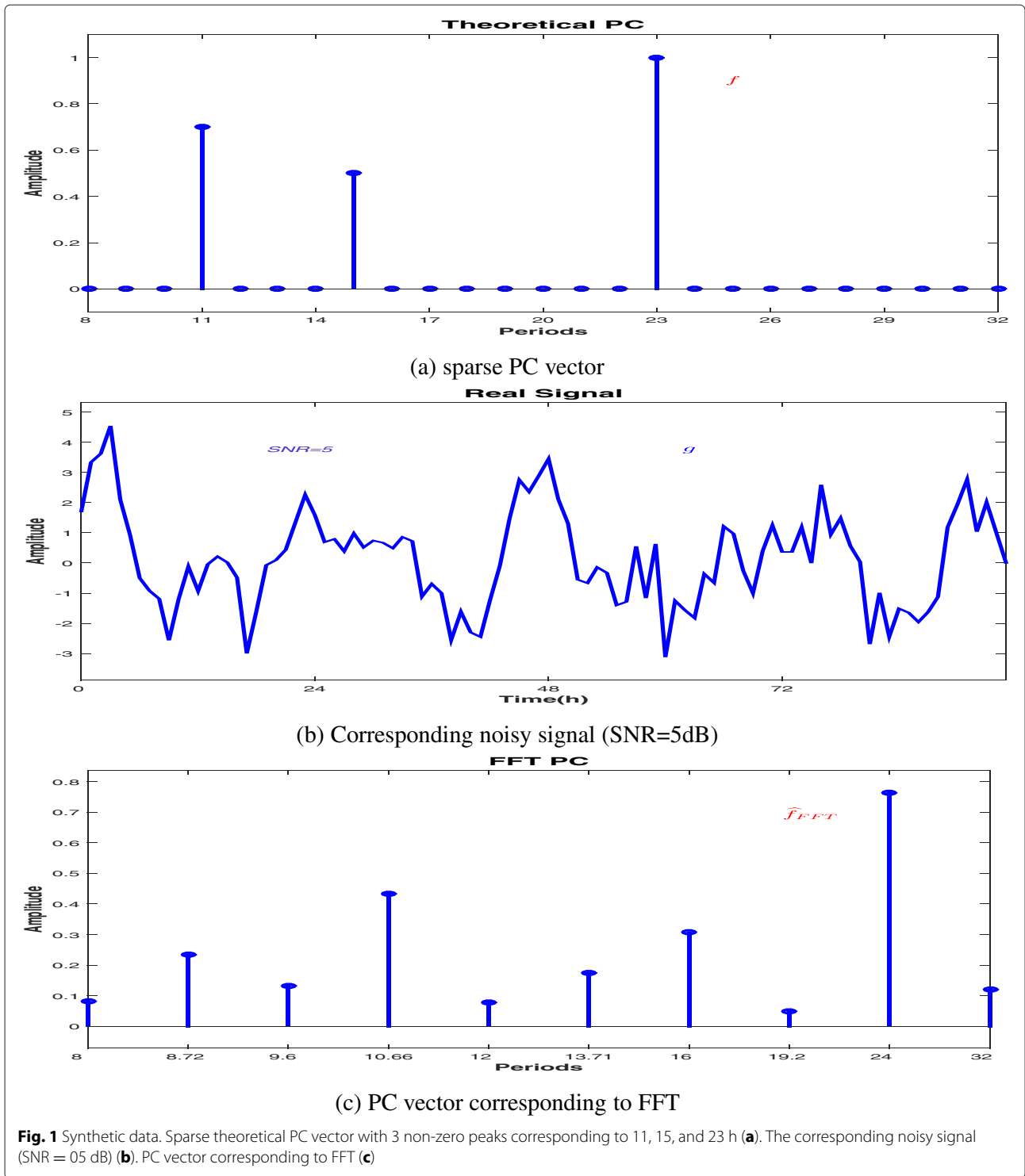
The article is positioned in the context of the need of a method capable to estimate the PC vector of a time series in the following conditions: (a) very limited number of data (4-day length) for circadian periodic components ( $24 \pm 6$  h) estimation and (b) precision that can be adjusted depending on the chronobiological context, 1-hour precision required in the particular experiment discussed in this article. The method proposed in this article formulates the estimation of the PC vector as an inverse problem, using the general Bayesian inference to

infer the unknowns of the considered linear model. This approach is presented in Section 3. A hierarchical prior model is considered, using the Student's  $t$  distribution (expressed as the marginal of a normal-inverse gamma bivariate distribution) as the sparsity enforcing prior law for the PC vector and assigning prior distributions for the hyperparameters involved in the model, namely the variances associated with the PC vector and the noise (Subsection 3.2 and Section 4). From the analytical expression of the joint posterior law of the unknown PC vector and hyperparameters, obtained via Bayes rule, the unknowns are estimated via joint maximum a posteriori (JMAP) (Subsection 4.1) or posterior mean (PM). For the PM estimator, the expression of the posterior law is approximated by a separable one, via the variational Bayesian approximation (VBA), using the Kullback-Leibler (KL) divergence. For the PM estimation, two possibilities are considered: an approximation with a partially separable law (Subsection 4.2) and one with a full separable one (Subsection 4.3). Simulation results on synthetic data (5 dB) and real data in cancer treatment applications are presented in Section 5. More simulations for the synthetic case (10 and 15 dB) are presented in the Additional file 1.

## 2 Classical Fourier transform methods

The spectral analysis for time series is a well-known subject in literature for a very long time. The most used methods are the FFT-based methods, which are widely used for many applications in signal processing community, having obvious advantages: the FFT-based methods are well known, well understood, and fast. Nevertheless, the particularities of the biomedical signals considered in chronobiology experiments show that the classical methods present certain limitations. In particular, for short time series relative to the dominant period (in the experiment considered in this article, a 96 h recorded signal relative to an  $\sim 24$  h dominant period, linked with the circadian clock) the precision given by the FFT methods is by far insufficient to determine the exact PC vector, since via the FFT-based methods, the frequency axis is linear, but as a function of the periods, it is not. In particular, for a 4-day (96 h) recorded signal, beside the 24 h corresponding periodic component, the nearest amplitudes in the PC vector correspond to 32 and 19 h (Fig. 1).

More general, if the prior knowledge sets the dominant period around a value  $P$  in order to obtain a PC vector that contains the period  $P$  and also the periods  $P - 1$  and  $P + 1$ , the signal must be observed for  $(P - 1)(P + 1)$  periods. In chronobiology applications, where the circadian period is around 24 h, a signal should be recorded for 575 days in order to obtain a periodic component vector that contains 23-, 24-, and 25-h periods.

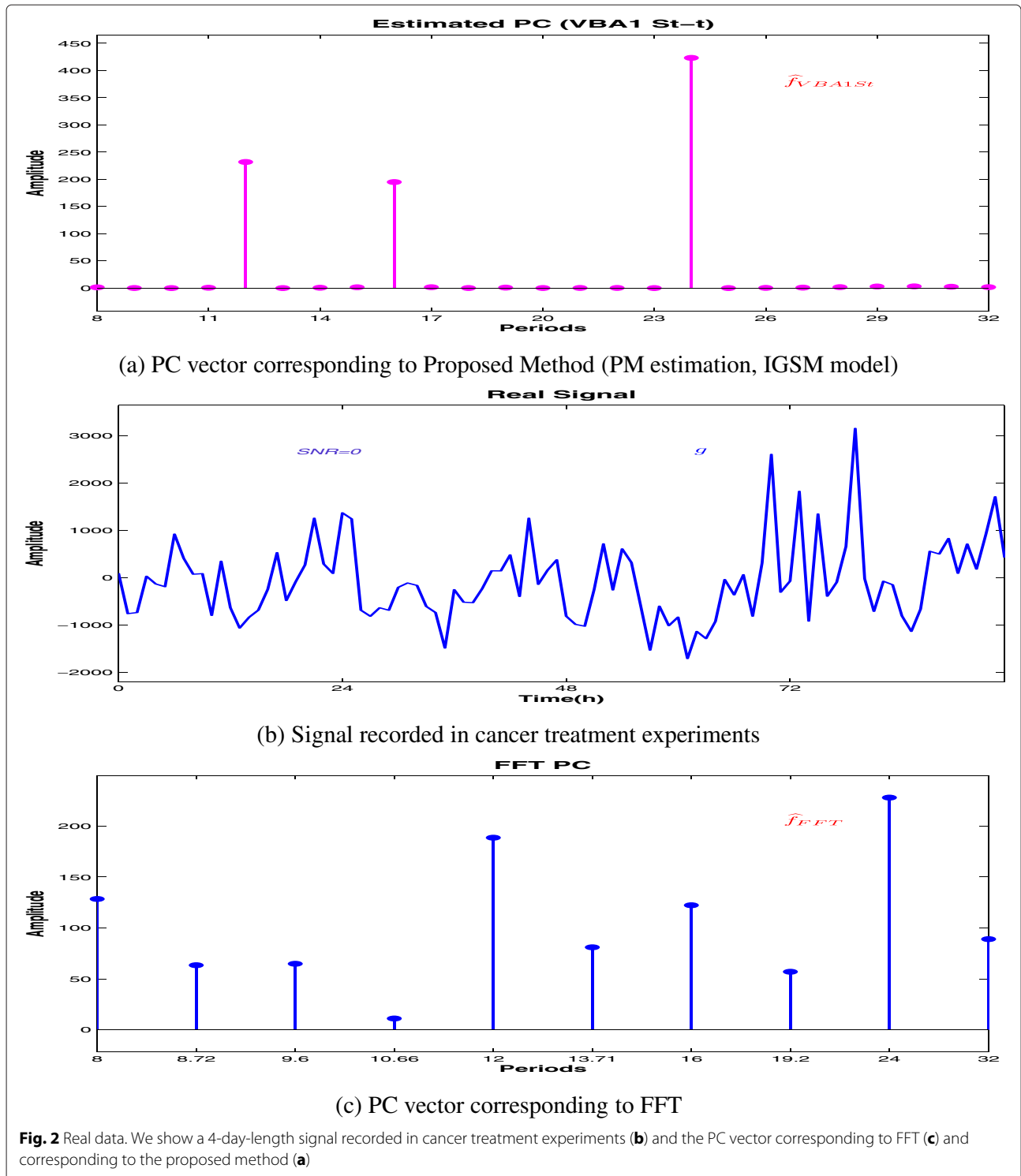


As an example, Fig. 1b presents a 4 day synthetic signal corresponding to a known PC vector (showed in Fig. 1a) and the corresponding PC vector obtained via FFT (Fig. 1c) (presented for the interval between 8 and 32 h, the circadian domain). In the synthetic PC vector,

the non-zero periods are set for 11, 15, and 23 h (dominant peak). The FFT estimates the dominant peak at 24 h, due to the fact that the time series observation period is limited to 96 h. In such conditions, it offers no information for the real positions, 11, 15, 23 h. It

also offers no informations for the peaks in the interval [20 : 31], except the estimation for 24 h. For similar signals corresponding to PC vectors having the principal peak around 24 h, the FFT will estimate the principal peak at 24 h. Another example is presented in Fig. 2b, a

4-day-length signal recorded in an experiment in chronobiology. The FFT PC vector presents peaks corresponding only to 8, 8.72, 9.6, 10.66, 12, 13.71, 16, 19.2, 24, and 32 h inside the considered interval [8 : 32]. The periods corresponding to 24, 12, and 8 h can be associated with



the presence of the circadian rhythm, expressed by the principal peak and the corresponding harmonics. But the presence of the other peaks can be more difficult to be interpreted by the biologists. Another drawback of FFT is the difficulty of selecting the peaks corresponding to the presence of a biological phenomena and peaks that are explained by error measures and uncertainties. This article is proposing a method that can estimate the PC vector, in the conditions (a) and (b) described in Section 1, taking into account the uncertainties and being able to distinguish between the peaks: the ones corresponding to the biological phenomena and the ones explained by other factors, producing a sparse PC vector (Fig. 2a). We think that this result will be more comprehensible by the biologists.

### 3 Inverse problem approach and general Bayesian inference

The first step in the proposed method for improving the precision consists in the inverse problem approach. We formulate the relation between the available data  $\mathbf{g}$  and the unknown PC vector  $\mathbf{f}$  as an inverse problem. The inverse Fourier transform provides a linear relation between the known biomedical signal  $\mathbf{g}$  and the PC vector  $\mathbf{f}$ . Using the elements of the matrix corresponding to the inverse Fourier transform, the linear relation is described by the following equation:

$$g(t_i) \simeq \sum_{j=1}^M f(p_j) e^{2\pi j \frac{1}{p_j} t_i}, \quad i \in \{1, \dots, N\}, \quad j \in \{1, \dots, M\}. \tag{1}$$

Introducing the notations  $g(t_i) = g_i$  and  $f(t_j) = f_j$ , Eq. (1) becomes

$$g_i \simeq \sum_{j=1}^M f_j e^{2\pi j \frac{1}{p_j} t_i}, \quad i \in \{1, \dots, N\} \quad j \in \{1, \dots, M\} \rightarrow \mathbf{g} \simeq \mathbf{H}\mathbf{f}. \tag{2}$$

Due to the potential modeling and measurement errors, we need to account for errors and uncertainties, so the linear model given by the inverse Fourier transform is completed by introducing the error vector  $\boldsymbol{\epsilon}$ , obtaining the forward model, Eq. (3):

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon}, \tag{3}$$

where we have used the following notations:

- $\mathbf{g}$  represents the observed data, i.e., the chronobiological time series:  
 $\mathbf{g} = [g_1, g_2 \dots g_N]^T \in \mathcal{M}_{N \times 1}$ , an  $N$ -dimensional vector

- $\mathbf{f}$  represents the unknowns, i.e., the PC vector:  
 $\mathbf{f} = [f_1, f_2, \dots, f_M]^T \in \mathcal{M}_{M \times 1}$ , a  $M$ -dimensional vector
- $\boldsymbol{\epsilon}$  represents the errors:  
 $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \dots, \epsilon_N]^T \in \mathcal{M}_{N \times 1}$ , is an  $N$ -dimensional vector

The goal is to estimate the unknowns of the model, Eq. (3), i.e., the PC vector  $\mathbf{f}$  and the error vector  $\boldsymbol{\epsilon}$ . In this paper, we propose an inversion based on general Bayesian inference, building a hierarchical model and estimating the unknowns from the posterior probability density function, using the available data  $\mathbf{g}$ . The estimated  $\hat{\mathbf{f}}$  and the corresponding estimated signal  $\hat{\mathbf{g}} = \mathbf{H}\hat{\mathbf{f}}$  are compared with  $\mathbf{f}$  (only in the synthetic case) and  $\mathbf{g}$ , using as a measure of performance Eq. (4).

$$\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}\|_2^2}{\|\mathbf{f}\|_2^2}; \quad \delta \mathbf{g} = \frac{\|\mathbf{g} - \hat{\mathbf{g}}\|_2^2}{\|\mathbf{g}\|_2^2}. \tag{4}$$

For the application considered in this paper, the matrix  $\mathbf{H}$  used in the model presented in Eq. (3) has very high conditioning numbers, so the problem is ill-conditioned. As mentioned above, in this paper, we focus on an inversion based on general Bayesian inference. Nevertheless, in literature, many other approaches are possible. One particular case of the considered linear model is the case where the error vector is neglected ( $\boldsymbol{\epsilon} = \mathbf{0}$ ), and the matrix  $\mathbf{H}$  is invertible and orthogonal, i.e.,  $\mathbf{H}^T \mathbf{H} = \mathbf{I}$ . This is the case of the FT matrix with  $M = N$ . Then, the solution is given by  $\hat{\mathbf{f}} = \mathbf{H}^T \mathbf{g}$ , which corresponds to IFT. However, in general as in our case  $M \neq N$ . When  $M < N$ , a minimum norm solution  $\hat{\mathbf{f}}_{\text{MN}} = (\mathbf{H}\mathbf{H}^T)^{-1} \mathbf{H}^T \mathbf{g}$  can be obtained, and when  $M > N$ , the classical solution is the least square solution  $\hat{\mathbf{f}}_{\text{LS}} = \mathbf{H}^T (\mathbf{H}\mathbf{H}^T)^{-1} \mathbf{g}$ . Since in the case of chronobiological times series the matrix  $\mathbf{H}$  is proved to have a very high conditioning number, those generalized inverse solutions are, in general, too sensitive to the errors due to the ill-conditioning of the matrix  $\mathbf{H}$ . The regularization methods can partially solve this difficulty. For example, the regularization methods such as truncated single value decomposition (TSVD) or Tikhonov regularization methods (TRM) can be used, but the solutions depend on the threshold in the first case (TSVD) and on the regularization parameter in the second case. When  $M \neq N$  and when the error vector is not neglected ( $\boldsymbol{\epsilon} \neq \mathbf{0}$ ), the regularization methods can still be applied and an estimation can be obtained for  $\mathbf{f}$  and  $\boldsymbol{\epsilon}$ , but with the following drawbacks: in general, determining the regularization parameters is difficult and there is not a good way to handle other a priori knowledge we may have on the noise statistics and on the unknowns.

### 3.1 Bayesian inference

A fundamental particularity of the proposed method is the use of the prior knowledge. In this article, we adopt a Bayesian approach. The Bayesian approach for times series was considered in [41–53]. However, the lack of data makes the proposed methods inefficient for our case. In Bayesian inference, the fundamental relation is given by the Bayes rule:

$$p(\mathbf{f}|\mathbf{g}, \theta_1, \theta_2) = \frac{p(\mathbf{g}|\mathbf{f}, \theta_1) p(\mathbf{f}|\theta_2)}{p(\mathbf{g}|\theta_1, \theta_2)}, \theta = (\theta_1, \theta_2), \quad (5)$$

where  $\theta$  represents the hyperparameters that appear in the model. In general, we may not know the hyperparameters  $\theta$ , and this is also our case. The hyperparameters represents the variances associated with the noise  $\epsilon$  and with  $\mathbf{f}$ , which are unknown. We need to estimate them, too. This can be done via the joint posterior law:

$$p(\mathbf{f}, \theta_1, \theta_2|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}, \theta_1) p(\mathbf{f}|\theta_2) p(\theta_1) p(\theta_2). \quad (6)$$

Such an extension presents two particular advantages: one advantage is evidently the possibility of estimating the hyperparameters and obtaining numerical values for variances and the second one is that such an approach can be developed into a non-supervised algorithm.

### 3.2 Hierarchical prior models

The hierarchical model represents the set of probability density functions assigned for the probabilities involved in (6), namely the assignment of the prior  $p(\mathbf{g}|\mathbf{f}, \theta_1)$ , the likelihood  $p(\mathbf{f}|\theta_2)$ , and the hyperparameters priors  $p(\theta_1)$ ,  $p(\theta_2)$ . The prior biological knowledge leads to the search of good sparsity enforcing priors. In literature [54], certain classes of distribution (heavy-tailed, mixture models) are well known as good sparsity enforcing priors. In this paper, we consider a general infinite Gaussian scale mixture (IGSM) hierarchical model [55]. The prior distribution for the PC vector is a Student's  $t$  distribution expressed via a normal-inverse gamma distribution. The error vector is also modeled using the IGSM, considering non-stationary variances for the noise, generalizing the results from [56]. In Section 5, during the simulations results, we include comparisons with the Gaussian hierarchical model for the synthetic data.

## 4 Hierarchical model infinite Gaussian scale mixture

In the first step, we model the error vector  $\epsilon$ . We propose to use a non-stationary Gaussian model:

$$p(\epsilon_i|\nu_{\epsilon_i}) = \mathcal{N}(\epsilon_i|0, \nu_{\epsilon_i}), \quad i \in \{1, 2, \dots, N\}, \quad (7)$$

where  $\nu_{\epsilon_i}$  are considered to be unknowns. For having the possibility to estimate them, we model them as inverse gamma distributions:

$$p(\nu_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \mathcal{IG}(\nu_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}), \quad i \in \{1, 2, \dots, N\}. \quad (8)$$

Doing this, we model the error vector as an infinite Gaussian scale mixture:

$$\begin{cases} p(\epsilon|\nu_{\epsilon}) = \mathcal{N}(\epsilon|\mathbf{0}, \mathbf{V}_{\epsilon}) \\ p(\nu_{\epsilon}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \prod_{i=1}^N \mathcal{IG}(\nu_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}), \end{cases} \quad (9)$$

where we introduced the vector  $\nu_{\epsilon}$  and the corresponding diagonal matrix  $\mathbf{V}_{\epsilon}$ :

$$\nu_{\epsilon} = [\nu_{\epsilon_1} \dots \nu_{\epsilon_i} \dots \nu_{\epsilon_N}]^T; \quad \mathbf{V}_{\epsilon} = \text{diag}[\nu_{\epsilon}]. \quad (10)$$

The likelihood  $p(\mathbf{g}|\mathbf{f}, \nu_{\epsilon})$  is obtained using the considered linear model, Eq. (3), and the assigned distribution for the error vector  $\epsilon$  conditioned by the variance  $\nu_{\epsilon}$ , Eq. (9). The distribution modeling the likelihood is also a multivariate normal distribution, having the same covariance matrix  $\mathbf{V}_{\epsilon}$  and the mean  $\mathbf{H}\mathbf{f}$ :

$$p(\mathbf{g}|\mathbf{f}, \nu_{\epsilon}) = \mathcal{N}(\mathbf{g}|\mathbf{H}\mathbf{f}, \mathbf{V}_{\epsilon}). \quad (11)$$

The proposed prior distribution is a Student's  $t$  distribution, in order to enforce the sparsity and use the prior knowledge of reduced number of periods in the PC vector. While a direct assignment of a Student's  $t$  distribution for the prior law  $p(\mathbf{f})$  leads to a non-quadratic criterion when estimating  $\mathbf{f}$ , the Student's  $t$  distribution corresponding to the prior law can be expressed as an infinite Gaussian scale mixture, modeling the inverse variance as a gamma distribution or the variance as an inverse gamma distribution. For the variance of  $\mathbf{f}$ , we assume a general model:

$$\nu_{\mathbf{f}} = [\nu_{f_1} \dots \nu_{f_j} \dots \nu_{f_M}]^T; \quad \mathbf{V}_{\mathbf{f}} = \text{diag}[\nu_{\mathbf{f}}]. \quad (12)$$

The prior law is then defined as an infinite Gaussian scale mixture, via  $\nu_{\mathbf{f}}$ :

$$\begin{cases} p(\mathbf{f}|\nu_{\mathbf{f}}) = \mathcal{N}(\mathbf{f}|\mathbf{0}, \mathbf{V}_{\mathbf{f}}) \\ p(\nu_{\mathbf{f}}|\alpha_{f_0}, \beta_{f_0}) = \prod_{j=1}^M \mathcal{IG}(\nu_{f_j}|\alpha_{f_0}, \beta_{f_0}). \end{cases} \quad (13)$$

The error variance priors, Eq. (9), the likelihood, Eq. (11), and the prior, Eq. (13), represents the IGSM hierarchical model. The analytical form is presented in Eq. (14):



$$\begin{cases} p(\mathbf{g}|\mathbf{f}, \mathbf{v}_\epsilon) = \mathcal{N}(\mathbf{g}|\mathbf{H}\mathbf{f}, \mathbf{V}_\epsilon) \propto \det^{-\frac{1}{2}}(\mathbf{V}_\epsilon) \exp\left\{-\frac{1}{2}\|\mathbf{V}_\epsilon^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f})\|^2\right\} \\ p(\mathbf{f}|\mathbf{v}_f) = \mathcal{N}(\mathbf{f}|\mathbf{0}, \mathbf{V}_f) \propto \det^{-\frac{1}{2}}(\mathbf{V}_f) \exp\left\{-\frac{1}{2}\|(\mathbf{V}_f)^{-\frac{1}{2}}\mathbf{f}\|^2\right\} \\ p(\mathbf{v}_\epsilon|\alpha_{\epsilon 0}, \beta_{\epsilon 0}) = \prod_{i=1}^N \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon 0}, \beta_{\epsilon 0}) \propto \prod_{i=1}^N v_{\epsilon_i}^{-(\alpha_{\epsilon 0}+1)} \exp\left\{-\sum_{i=1}^N \beta_{\epsilon 0} v_{\epsilon_i}^{-1}\right\} \\ p(\mathbf{v}_f|\alpha_{f 0}, \beta_{f 0}) = \prod_{j=1}^M \mathcal{IG}(v_{f_j}|\alpha_{f 0}, \beta_{f 0}) \propto \prod_{j=1}^M v_{f_j}^{-(\alpha_{f 0}+1)} \exp\left\{-\sum_{j=1}^M \beta_{f 0} v_{f_j}^{-1}\right\}. \end{cases} \quad (14)$$

From the hierarchical model, the posterior distribution can be obtained via the proportionality relation considered in Eq. (6):

$$\hat{v}_{f_j \text{JMAP}} = \frac{\beta_{f 0} + \frac{f_j^2}{2}}{\alpha_{f 0} + 1 + \frac{1}{2}}, \quad (17c)$$

$$p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}, \mathbf{v}_\epsilon) p(\mathbf{f}|\mathbf{v}_f) p(\mathbf{v}_\epsilon|\alpha_{\epsilon 0}, \beta_{\epsilon 0}) p(\mathbf{v}_f|\alpha_{f 0}, \beta_{f 0}). \quad (15)$$

**4.1 Joint MAP estimation**

The joint maximum a posteriori, a point estimator of the unobserved quantities  $\mathbf{f}$ ,  $\mathbf{v}_\epsilon$ ,  $\mathbf{v}_f$  on the basis of the available data  $\mathbf{g}$  is defined as:

$$(\hat{\mathbf{f}}, \hat{\mathbf{v}}_\epsilon, \hat{\mathbf{v}}_f) = \arg \max_{(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f)} p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g}) = \arg \min_{(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f)} \mathcal{L}(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f), \quad (16)$$

where for the second equality, we have defined the criterion  $\mathcal{L}(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f) = -\ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g})$ . The MAP estimator is the solution minimizing the criterion  $\mathcal{L}(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f)$ . This can be done via alternate optimization with respect to each of the unknowns. The computation details are presented in Appendix 1. Here, we present the final results in Eqs. (17a), (17b), and (17c).

$$\hat{\mathbf{f}}_{\text{JMAP}} = [\mathbf{H}^T \mathbf{V}_\epsilon^{-1} \mathbf{H} + \mathbf{V}_f^{-1}]^{-1} \mathbf{H}^T \mathbf{V}_\epsilon^{-1} \mathbf{g} \quad (17a)$$

$$\hat{v}_{\epsilon_i \text{JMAP}} = \frac{\beta_{\epsilon 0} + \frac{1}{2}(g_i - \mathbf{H}_i \hat{\mathbf{f}})^2}{\alpha_{\epsilon 0} + 1 + \frac{1}{2}} \quad (17b)$$

where  $\mathbf{H}_i$  represents the line  $i$  from the matrix  $\mathbf{H}$ . The iterative algorithm obtained via JMAP estimation is presented in Fig. 3. The algorithm is compared with the one corresponding to the posterior mean estimation in Section 5.

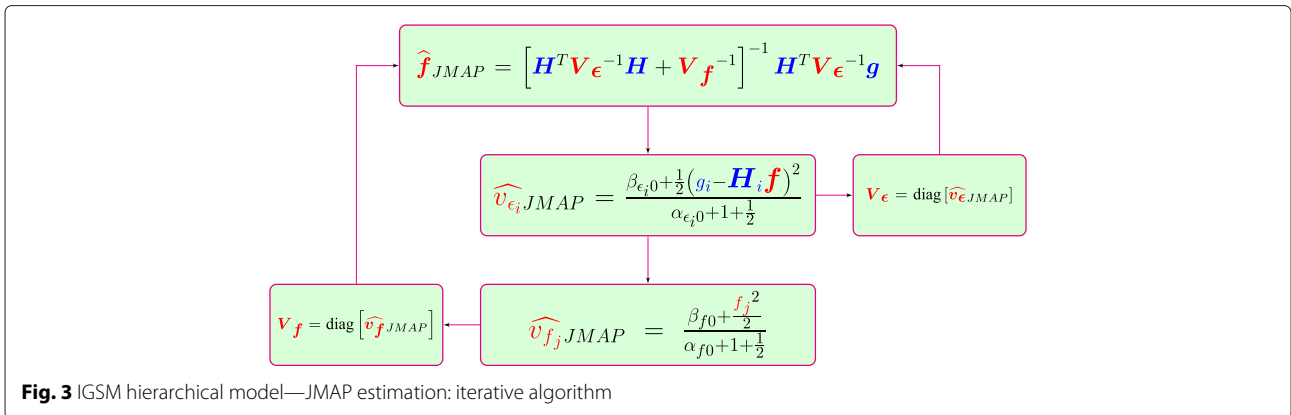
**4.2 Posterior mean (via VBA) IGSM (partial separability)**

The posterior mean estimates the mean of the posterior distribution. One of the advantages of this estimator is the fact that it minimizes the mean square error (MSE). In particular, the posterior distribution obtained from the considered hierarchical model is not a separable distribution, making the analytical computation of PM very difficult. One way to compute the PM in this case is to first approximate the posterior law  $p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g})$  with a separable law  $q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g})$  [54]:

$$p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g}) \approx q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g}) = q_1(\mathbf{f}) q_2(\mathbf{v}_\epsilon) q_3(\mathbf{v}_f), \quad (18)$$

where we have used the notations

$$q_2(\mathbf{v}_\epsilon) = \prod_{i=1}^N q_{2i}(v_{\epsilon_i}) ; q_3(\mathbf{v}_f) = \prod_{j=1}^M q_{3j}(v_{f_j}), \quad (19)$$



**Fig. 3** IGSM hierarchical model—JMAP estimation: iterative algorithm

The approximate  $q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  is obtained by minimizing the Kullback-Leibler divergence, defined as:

$$\begin{aligned} & \text{KL}(q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) : p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})) \\ &= \iint \dots \int q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \ln \frac{q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})}{p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})} d\mathbf{f} d\mathbf{v}_\epsilon d\mathbf{v}_f, \end{aligned} \quad (20)$$

where we also used the notations:

$$d\mathbf{v}_\epsilon = \prod_{i=1}^N dv_{\epsilon_i} ; \quad d\mathbf{v}_f = \prod_{j=1}^M dv_{f_j}, \quad (21)$$

Like in the MAP case, the minimization can be done via alternate optimization resulting in the following proportionalities:

$$\begin{cases} q_1(\mathbf{f}) \propto \exp \left\{ \left( \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \right)_{q_2(\mathbf{v}_\epsilon) q_3(\mathbf{v}_f)} \right\} \\ q_{2i}(\mathbf{v}_{\epsilon_i}) \propto \exp \left\{ \left( \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \right)_{q_1(\mathbf{f}) q_{2-i}(\mathbf{v}_{\epsilon_i}) q_3(\mathbf{v}_f)} \right\}, i \in \{1, 2, \dots, N\} \\ q_{3j}(\mathbf{v}_{f_j}) \propto \exp \left\{ \left( \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \right)_{q_1(\mathbf{f}) q_2(\mathbf{v}_\epsilon) q_{3-j}(\mathbf{v}_{f_j})} \right\}, j \in \{1, 2, \dots, M\} \end{cases} \quad (22)$$

where we used the notations:

$$\begin{aligned} q_{2-i}(\mathbf{v}_{\epsilon_i}) &= \prod_{k=1, k \neq i}^N q_{2k}(\mathbf{v}_{\epsilon_k}) ; \quad q_{3-j}(\mathbf{v}_{f_j}) \\ &= \prod_{k=1, k \neq j}^M q_{3k}(\mathbf{v}_{f_k}) ; \quad \langle u(x) \rangle_{v(y)} = \int u(x)v(y) dy. \end{aligned} \quad (23)$$

From the proportionalities showed in Eq. (22), we derive the probability distributions corresponding to  $q_1(\mathbf{f})$ ,  $q_{2i}(\mathbf{v}_{\epsilon_i})$ ,  $q_{3j}(\mathbf{v}_{f_j})$  and the corresponding parameters. The detailed computations are presented in Appendix 2. Here, we only present the general strategy: in the first step, developing the proportionality corresponding to  $q_1(\mathbf{f})$ , we obtain an expression of an exponential, having as argument a quadratic criterion, leading to the conclusion that  $q_1(\mathbf{f})$  is a multivariate normal distribution. By minimizing the criterion, we obtain the analytical expression of the corresponding mean. The variance is obtained by identification. However, in this stage, both the analytical expressions of the mean and variance depend on expectancies corresponding to the two variances involved in the model, i.e.,  $\mathbf{v}_{\epsilon_i}$  and  $\mathbf{v}_{f_j}$ . In the second step, developing the proportionalities corresponding to  $q_{2i}(\mathbf{v}_{\epsilon_i})$  and  $q_{3j}(\mathbf{v}_{f_j})$ , we establish that they are both inverse gamma distributions. This is done using that the expectancies containing  $\mathbf{f}$  can be handled because  $q_1(\mathbf{f})$  was proved to be a multivariate Normal distribution in the previous step. Then, using the fact that  $q_{2i}(\mathbf{v}_{\epsilon_i})$  and  $q_{3j}(\mathbf{v}_{f_j})$  are inverse gamma distributions, the expectancies that appear in the expressions of

the mean and variance corresponding to the multivariate normal distribution  $q_1(\mathbf{f})$  can be computed. We establish analytical expressions for all the parameters of the distributions. The analytical expressions of the parameters are presented in the Eqs. (24a), (24b), and (24c).

$$q_1(\mathbf{f}) = \mathcal{N}(\mathbf{f} | \hat{\mathbf{f}}_{\text{PM}}, \hat{\Sigma}), \quad \begin{cases} \hat{\mathbf{f}}_{\text{PM}} = (\mathbf{H}^T \widehat{\mathbf{V}}_\epsilon^{-1} \mathbf{H} + \widehat{\mathbf{V}}_f^{-1})^{-1} \mathbf{H}^T \widehat{\mathbf{V}}_\epsilon^{-1} \mathbf{g} \\ \hat{\Sigma} = (\mathbf{H}^T \widehat{\mathbf{V}}_\epsilon^{-1} \mathbf{H} + \widehat{\mathbf{V}}_f^{-1})^{-1} \end{cases} \quad (24a)$$

$$q_{2i}(\mathbf{v}_{\epsilon_i}) = \mathcal{IG}(\mathbf{v}_{\epsilon_i} | \alpha_{\epsilon_i}, \beta_{\epsilon_i}), \quad \begin{cases} \alpha_{\epsilon_i} = \alpha_{\epsilon 0} + \frac{1}{2} \\ \beta_{\epsilon_i} = \beta_{\epsilon 0} + \frac{1}{2} \left[ \mathbf{H}_i \hat{\Sigma} \mathbf{H}_i^T + (\mathbf{g}_i - \mathbf{H}_i \hat{\mathbf{f}}_{\text{PM}})^2 \right] \end{cases} \quad (24b)$$

$$q_{3j}(\mathbf{v}_{f_j}) = \mathcal{IG}(\mathbf{v}_{f_j} | \alpha_{f_j}, \beta_{f_j}), \quad \begin{cases} \alpha_{f_j} = \alpha_{f 0} + \frac{1}{2} \\ \beta_{f_j} = \beta_{f 0} + \frac{1}{2} (\hat{f}_{\text{PM}j}^2 + \hat{\Sigma}_{jj}) \end{cases} \quad (24c)$$

Equation 24a provides the dependency of the parameters corresponding to the multivariate normal distribution  $q_1(\mathbf{f})$  and the others hyperparameters involved in the hierarchical model: the mean  $\hat{\mathbf{f}}_{\text{PM}}$  and the covariance matrix  $\hat{\Sigma}$  depend on  $\widehat{\mathbf{V}}_\epsilon^{-1}$  and  $\widehat{\mathbf{V}}_f^{-1}$ . Eq. (70) (in Appendix 2) defines  $\widehat{\mathbf{V}}_\epsilon^{-1}$  and  $\widehat{\mathbf{V}}_f^{-1}$  via  $\{\alpha_{\epsilon_i}, \beta_{\epsilon_i}\}, i \in \{1, 2, \dots, N\}$  and  $\{\alpha_{f_j}, \beta_{f_j}\}, j \in \{1, 2, \dots, M\}$ . For the mean and the variance, we obtain the following dependency:

$$\left\{ \alpha_{f_j}, \beta_{f_j} \right\}, \left\{ \alpha_{\epsilon_j}, \beta_{\epsilon_j} \right\} \longrightarrow \left[ \hat{\mathbf{f}}_{\text{PM}}, \hat{\Sigma} \right] \quad (25)$$

Equation 24b leads to the following dependency scheme:

$$\left[ \hat{\mathbf{f}}_{\text{PM}}, \hat{\Sigma} \right] \longrightarrow \left\{ \alpha_{\epsilon_j}, \beta_{\epsilon_j} \right\} \quad (26)$$

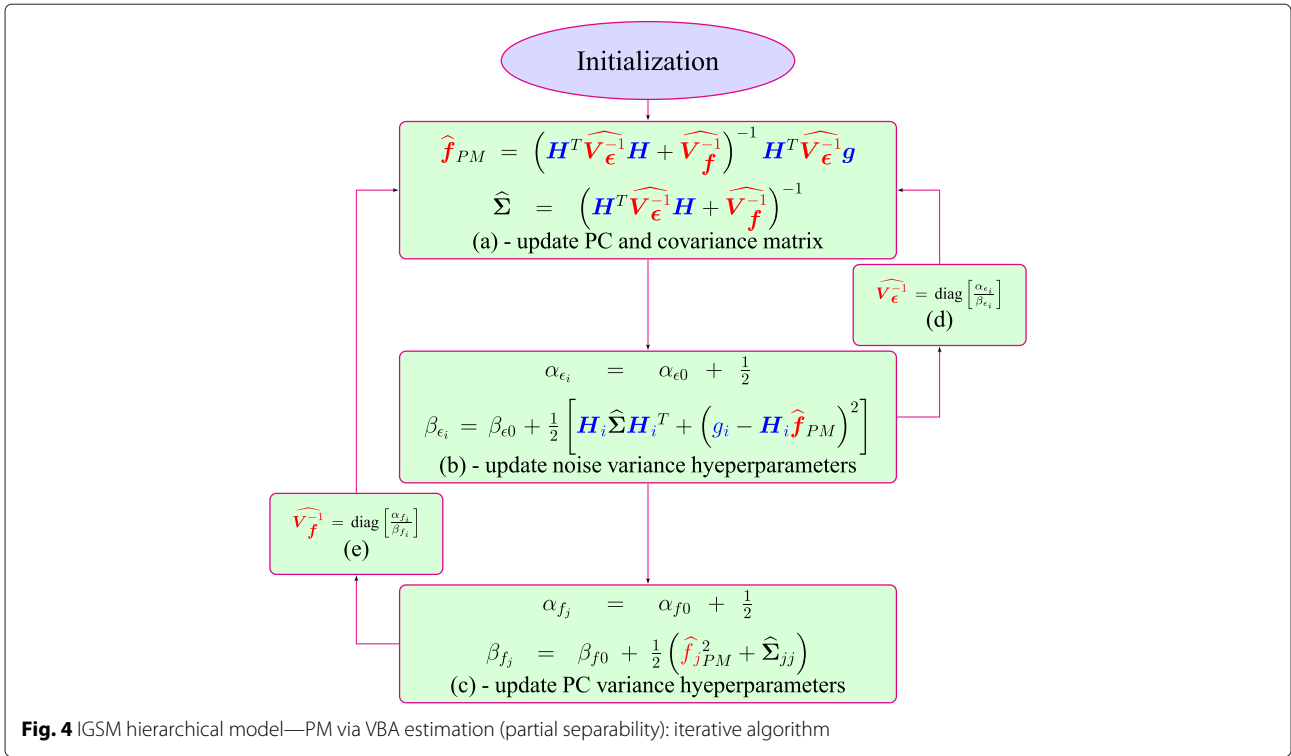
Equation 24c leads to the following dependency scheme:

$$\left[ \hat{\mathbf{f}}_{\text{PM}}, \hat{\Sigma} \right] \longrightarrow \left\{ \alpha_{f_j}, \beta_{f_j} \right\} \quad (27)$$

The dependencies presented in (25), (26), and (27) leads to an iterative algorithm with a parameter update for every step. The algorithm is as follows:

- Initialization
- Use Eqs. (24a) and (70) to compute  $\hat{\mathbf{f}}_{\text{PM}}, \hat{\Sigma}$
- Use Eq. (24b) to compute  $\{\alpha_{\epsilon_j}, \beta_{\epsilon_j}\}$  and  $\widehat{\mathbf{V}}_\epsilon^{-1}$
- Use Eq. (24c) to compute  $\{\alpha_{f_j}, \beta_{f_j}\}$  and  $\widehat{\mathbf{V}}_f^{-1}$

The iterative algorithm obtained using the PM estimation, via VBA partial separability, is presented in Fig. 4.



In order to initialize the algorithm, we define the matrices  $\widehat{V}_\epsilon^{-1(0)}$  and  $\widehat{V}_f^{-1(0)}$ , corresponding to the iteration zero of the algorithm. For the first iteration, using those values of matrices, the algorithm updates the estimations corresponding to the PC vector and the corresponding covariance matrix (a). Except the two matrices used, the other terms involved in the equations are known: the recorded signal  $\mathbf{g}$  and the matrix  $\mathbf{H}$ . After the PC vector and the covariance matrix are updated, they are used as terms in the equations updating the hyperparameters involved in the model. For updating the hyperparameters corresponding to the noise variances (b) and PC variances (c), the algorithm is using the estimation of the PC vector and the covariance matrix corresponding to the first iteration, obtained in (a). Then, the estimation corresponding to the noise variance (b) are used as input in (a), corresponding to the second iteration, via (d) and (e).

For initializing the algorithm, one of the possible choices is assigning values for the following parameters:  $\{\alpha_{f_j}^{(0)}, \beta_{f_j}^{(0)}\}$ ,  $j \in \{1, 2, \dots, M\}$  representing  $\widehat{V}_f^{-1(0)}$  and  $\{\alpha_{\epsilon_i}^{(0)}, \beta_{\epsilon_i}^{(0)}\}$ ,  $i \in \{1, 2, \dots, N\}$  representing  $\widehat{V}_\epsilon^{-1(0)}$ , corresponding to the step zero of the algorithm. This choice for the initialization procedure is sufficient, in the sense that the considered parameters from above represent all the necessary informations for starting the first iteration of the algorithm and computing all other parameters of the

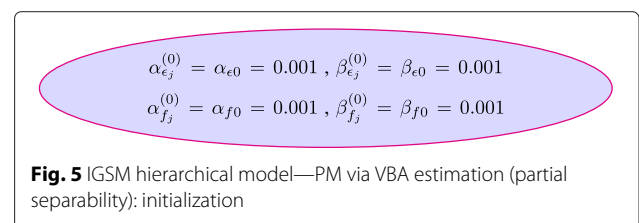
algorithm corresponding to step zero, i.e.,  $\widehat{\mathbf{f}}_{PM}^{(0)}$  and  $\widehat{\Sigma}^{(0)}$ . For the parameters  $\alpha_{\epsilon_i}^{(0)}$ ,  $\beta_{\epsilon_i}^{(0)}$  and  $\alpha_{f_j}^{(0)}$ ,  $\beta_{f_j}^{(0)}$ , we consider the following initialization:

$$\alpha_{\epsilon_j}^{(0)} = \alpha_{\epsilon 0} \quad , \quad \beta_{\epsilon_j}^{(0)} = \beta_{\epsilon 0} \quad , \quad \alpha_{f_j}^{(0)} = \alpha_{f 0} \quad , \quad \beta_{f_j}^{(0)} = \beta_{f 0} \quad (28)$$

A natural choice in this case is Non-Informative Prior Law (NIPL). The inverse gamma distribution is weak for parameters  $\alpha \rightarrow 0$  and  $\beta \rightarrow 0$ , so one possible choice is  $\alpha_{\epsilon 0} = \beta_{\epsilon 0} = 0.001$  and  $\alpha_{f 0} = \beta_{f 0} = 0.001$ . In particular, such an approach is consistent with a non-supervised algorithm. The considered initialization is presented in Fig. 5.

### 4.3 Posterior mean (via VBA) IGSM (full separability)

In Subsection 4.2, the Student's  $t$  model is considered and the PM estimator is used. The posterior law is approximated by a separable one, Eq. (18), where the notations



for  $q_2(\mathbf{v}_\epsilon)$  and  $q_3(\mathbf{v}_f)$ , introduced in Eq. (19), represent a full separability relative to  $\mathbf{v}_\epsilon$  and  $\mathbf{v}_f$ . But the prior law  $p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  is not approximated by a fully separable one since for  $q_1(\mathbf{f})$ , we consider a multivariate law modeling the vector  $\mathbf{f}$ . In this subsection, we investigate the development of the proposed model and the same PM estimator, but the posterior law is approximated by a fully separable law relative to all the unknowns, i.e., also for  $\mathbf{f}$ . The interest of such development concerns the applications where the precision required is high, making the numerical computations very costly. In this case, the posterior law from the hierarchical model  $p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  is approximated by a fully separable probability density function:

$$p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \approx q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) = \prod_{j=1}^M q_{1j}(f_j) \prod_{i=1}^N q_{2i}(v_{\epsilon_i}) \prod_{j=1}^M q_{3j}(v_{f_j}) = q_1(\mathbf{f}) q_2(\mathbf{v}_\epsilon) q_3(\mathbf{v}_f) \tag{29}$$

where we used the notation introduced in (19) and also the following notations, considered during this paragraph:

$$q_1(\mathbf{f}) = \prod_{j=1}^M q_{1j}(f_j) ; d\mathbf{f} = \prod_{j=1}^M dq_{1j}(f_j) ; q_{1-j}(f_j) = \prod_{k=1, k \neq j}^M q_{1k}(f_k) \tag{30}$$

Like in Subsection 4.2, the law  $q(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  is obtained by minimizing the Kullback-Leibler divergence, Eq. (20), via alternate optimization, obtaining the proportionalities presented in Eq. (31):

$$\begin{cases} q_{1j}(f_j) \propto \exp \left\{ \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \Big|_{q_{1-j}(f_j) q_2(\mathbf{v}_\epsilon) q_3(\mathbf{v}_f)} \right\} & j \in \{1, 2, \dots, M\} \\ q_{2i}(v_{\epsilon_i}) \propto \exp \left\{ \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \Big|_{q_1(\mathbf{f}) q_{2-i}(v_{\epsilon_i}) q_3(\mathbf{v}_f)} \right\}, & i \in \{1, 2, \dots, N\} \\ q_{3j}(v_{f_j}) \propto \exp \left\{ \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \Big|_{q_1(\mathbf{f}) q_2(\mathbf{v}_\epsilon) q_{3-j}(v_{f_j})} \right\}, & j \in \{1, 2, \dots, M\} \end{cases} \tag{31}$$

The detailed computations are presented in Appendix 3. The analytical expressions of the parameters are presented in Eqs. (32a), (32b), and (32c).

$$q_1(f_j) = \mathcal{N} \left( f_j | \widehat{f}_{j\text{PM}}, \text{var}_j \right), \begin{cases} \widehat{f}_{j\text{PM}} = \frac{\mathbf{H}^T \mathbf{V}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}^{-j} \mathbf{f}^{-j})}{\left\| (\mathbf{V}_\epsilon^{-1})^{1/2} \mathbf{H}^j \right\|^2 + v_{f_j}^{-1}} \\ \text{var}_j = \frac{1}{\left\| (\mathbf{V}_\epsilon^{-1})^{1/2} \mathbf{H}^j \right\|^2 + v_{f_j}^{-1}} \end{cases} \tag{32a}$$

$$q_{2i}(v_{\epsilon_i}) = \mathcal{IG} \left( v_{\epsilon_i} | \alpha_{\epsilon_i}, \beta_{\epsilon_i} \right), \begin{cases} \alpha_{\epsilon_i} = \alpha_{\epsilon_0} + \frac{1}{2} \\ \beta_{\epsilon_i} = \beta_{\epsilon_0} + \frac{1}{2} \left( g_i - \mathbf{H}_i \widehat{f}_{\text{PM}} \right) \end{cases} \tag{32b}$$

$$q_{3j}(v_{f_j}) = \mathcal{IG} \left( v_{f_j} | \alpha_{f_j}, \beta_{f_j} \right), \begin{cases} \alpha_{f_j} = \alpha_{f_0} + \frac{1}{2} \\ \beta_{f_j} = \beta_{f_0} + \frac{1}{2} \left( \widehat{f}_{j\text{PM}}^2 + \text{var}_j \right), \end{cases} \tag{32c}$$

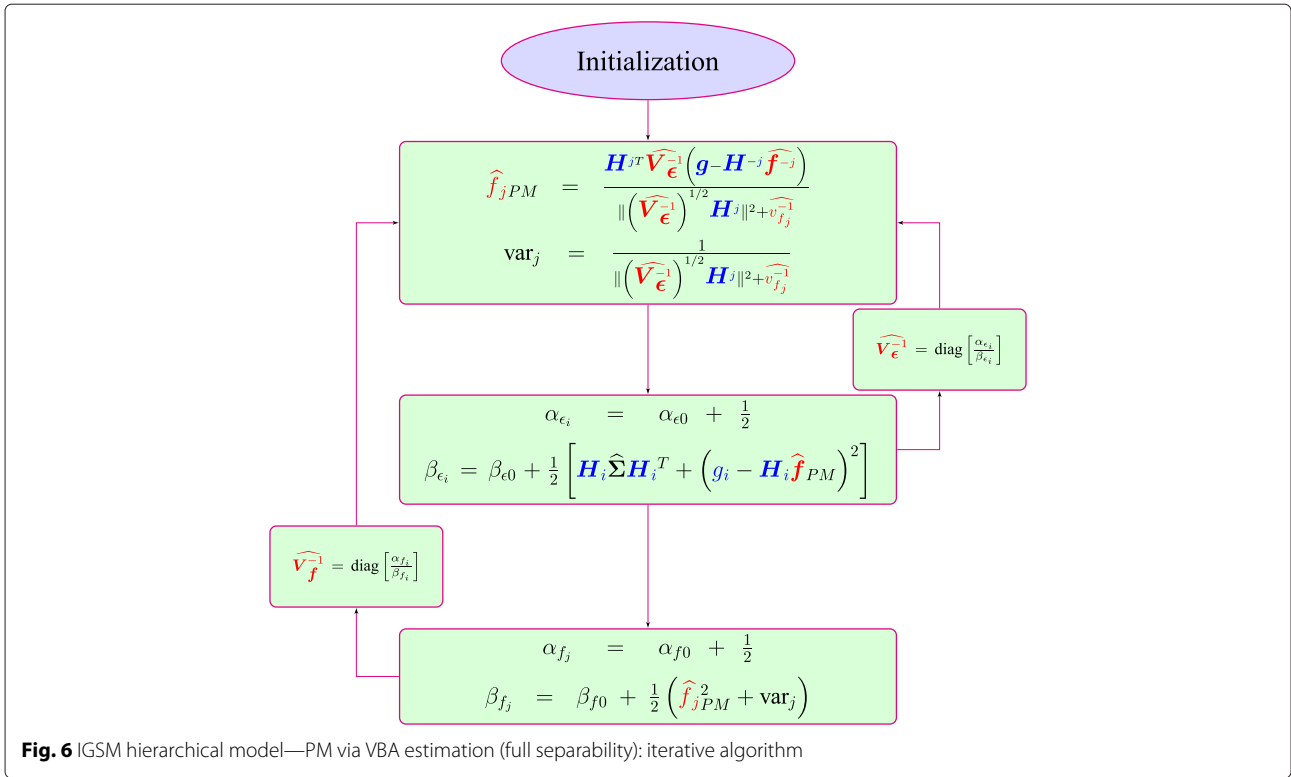
where  $\mathbf{H}^j$  represents the column  $j$  of the matrix  $\mathbf{H}$ ,  $\mathbf{H}^{-j}$  represents the matrix  $\mathbf{H}$  except the column  $j$ , and  $\mathbf{f}^{-j}$  represents the vector  $\mathbf{f}$  except the element  $f_j$ . The iterative algorithm is presented in Fig. 6. The initialization is done in the same conditions as in the partial separability case (Fig. 5).

### 5 Simulations

This section presents the simulations corresponding to synthetic and real data. For synthetic data, we compare five algorithms: joint MAP with Gaussian prior, posterior mean with Gaussian prior, joint MAP with IGSM prior, posterior mean (via VBA) with IGSM prior (partial separability), and posterior mean (via VBA) with IGSM prior (full separability). For each iterative algorithm, we present a comparison between the algorithm's estimation and the synthetic data, i.e., a comparison between  $\widehat{\mathbf{f}}_{\text{Method}}$  and  $\mathbf{f}$ , between  $\widehat{\mathbf{g}}_{\text{Method}}$  and  $\mathbf{g}$  and between  $\widehat{\mathbf{g}}_{\text{Method}}$  and  $\mathbf{g}_0$  theoretical signal ( $\mathbf{g}$  without noise). For every algorithm considered, we present the convergence analysis of the parameters and hyperparameters involved. Then, we present a comparison between the estimations of proposed algorithms and the classical FFT method. Finally, the proposed algorithms are tested 10 times over the same data, but different noise realization, in order to obtain the  $L_2$  error vector (the normalized difference between data and estimated data, considered for  $\mathbf{f}$ ,  $\mathbf{g}$ , and theoretical signal  $\mathbf{g}_0$ ) and compare the performances of each algorithm. These comparisons between error vectors corresponding to each algorithm are presented at the end of the subsection. For the synthetic data, we consider the following protocol: we consider a theoretical PC vector  $\mathbf{f}$  and the corresponding theoretical signal  $\mathbf{H}\mathbf{f}$  and we consider the corresponding signal  $\mathbf{g} = \mathbf{H}\mathbf{f} + \epsilon$ , by adding noise over the theoretical signal. In this article, we consider for the synthetic case three different levels of noise: 15, 10, and 5 dB. In this section, we include only the detailed simulations for the 5-dB case. The other two cases are presented in the Additional file 1. The considered signal represents a 4-day signal, sampled every hour. The matrix  $\mathbf{H}$  considered in this set of simulations is a cosine plus sine matrix.

#### 5.1 Synthetic data 05 dB

For testing, we have considered a 4-day signal, corresponding to a sparse PC vector, having non-zero values for 11, 15, and 23 h. We consider this particular structure for the following reason: we want to verify if the proposed



method can precisely distinguish the peaks inside the circadian domain. As we have mentioned, for such signals, via the FFT, we obtain a high peak corresponding to 24 h and the corresponding harmonics, but this method offers no information for certain values in the circadian domain. We have showed in Section 2 (Fig. 1) that a dominant period, corresponding to 23 h, is wrongly estimated at 24 h via FFT method and offers no other informations in the interval [20–31].

**5.1.1 Data 05 dB**

The PC vector  $f$ , theoretical signal  $g_0$ , and the signal  $g$  are presented in Fig. 7.

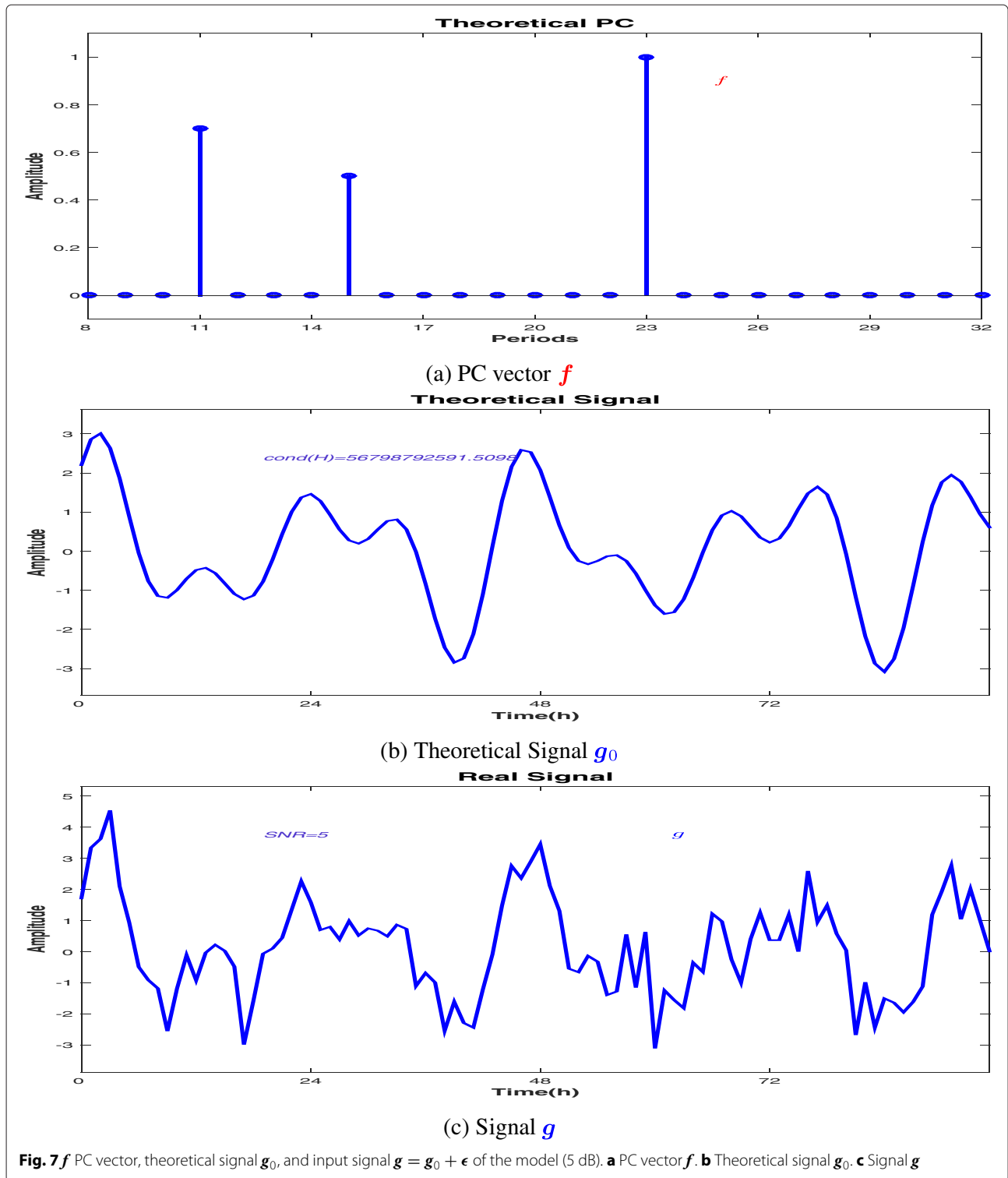
Figure 7a shows the theoretical PC, having the non-zero periods corresponding to 11, 15, and 23 h. All the other values in the PC vector are zero. Figure 7b presents the signal corresponding to the linear model considered in Eq. (3), neglecting the errors,  $g_0 = Hf$ . We note that the conditioning number of the matrix  $H$  is  $cond(H) = 56,798,792,591$ . All the simulations are done using the input as the noisy signal  $g$  corresponding to the linear model, Eq. (3), presented in Fig. 7c. We compare the estimated PC vector with the theoretical one (Fig. 7a) and the corresponding reconstructed signal with  $g_0$  and  $g$ . The comparison with the theoretical signal  $g_0$  is important in order to verify if the propose algorithm can distinguish the peaks corresponding to the biological phenomena from the ones corresponding to the noise.

**5.1.2 JMAP IGSM 05 dB**

A comparison between the synthetic data and the JMAP estimation, corresponding to the IGSM prior hierarchical model is presented in Fig. 8. We compare the theoretical PC vector  $f$  and the JMAP estimation  $\hat{f}_{JMAP}$ . We also present the comparison between the estimated  $\hat{g}_{JMAP}$  and  $g$  and the comparison between the estimated  $\hat{g}_{JMAP}$  and the theoretical signal (without noise)  $g_0$ .

The proposed method is searching for a sparse solution corresponding to the linear model, Eq. (3). The comparison between the theoretical signal  $g_0$  and  $\hat{g}_{JMAP}$  (Fig. 8b) shows that the proposed algorithm is converging to a solution that leads to a fairly accurate reconstruction, having the  $L_2$  norm error  $\delta g_0 = \frac{\|g_0 - \hat{g}_{JMAP}\|_2^2}{\|g_0\|_2^2} = 0.0524$ . For the PC vector, the reconstruction error is  $\delta f = \frac{\|f - \hat{f}_{JMAP}\|_2^2}{\|f\|_2^2} = 0.0726$ . For the JMAP estimation, the condition imposed for the searched solution, i.e., the sparsity is not respected (Fig. 8a). In fact, the alternate optimization algorithm considered for searching the JMAP solution is converging to a local minimum and the estimation errors corresponding to the JMAP estimation might be far from the example presented.

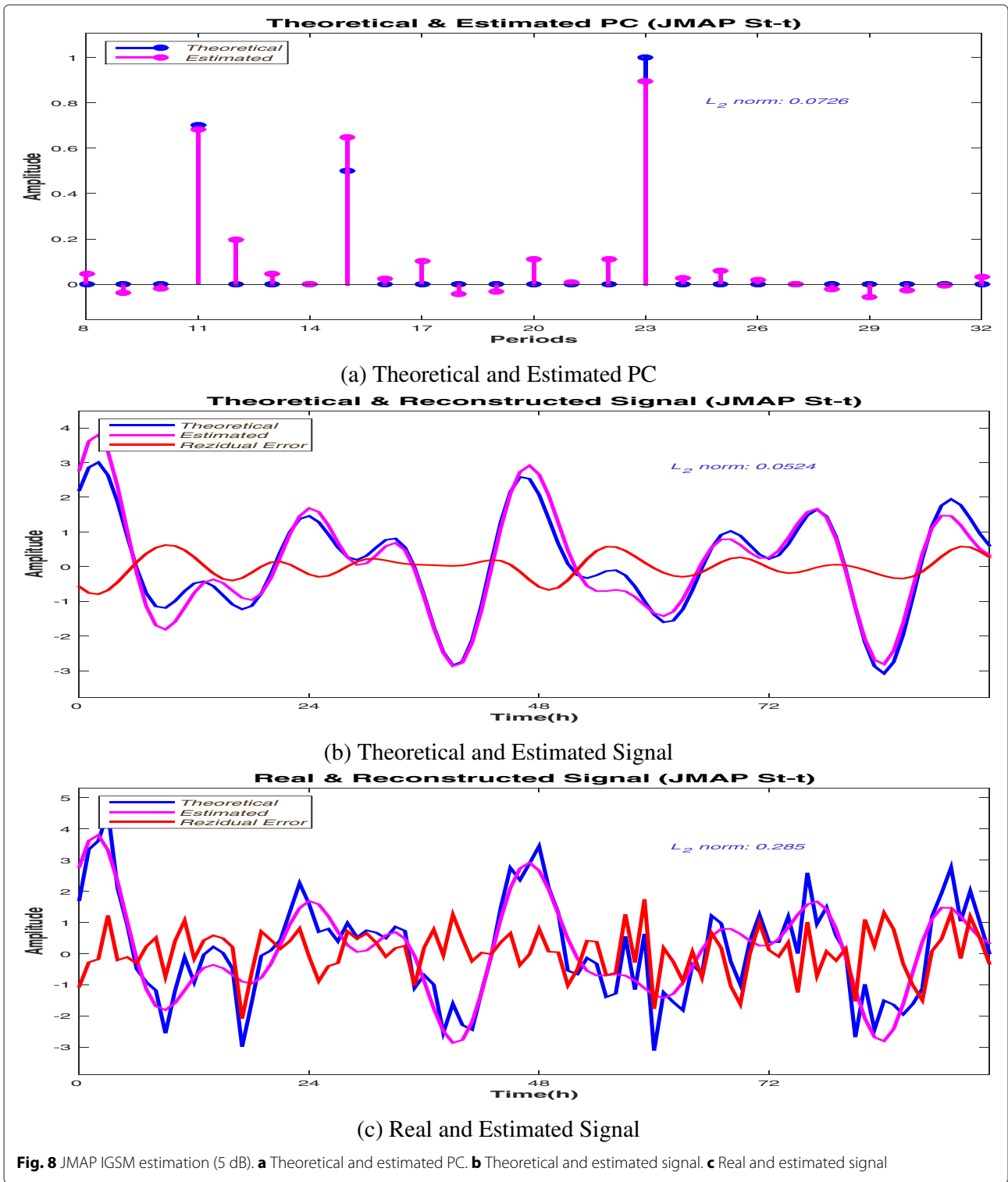
Figure 9a presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realization. As mentioned, the JMAP solution given by the alternate optimization algorithm is converging to a local minimum and the



estimation may be very inaccurate. We note that the figure presents a variation of  $L_2$  PC vector error reconstruction from 0.0524 to 4.2841. Important variations corresponding to the  $L_2$  error reconstruction for the theoretical signal  $g_0$  and signal  $g$  are presented in Fig. 9b, c.

### 5.1.3 PM (via VBA, partial separability) IGSM 05 dB

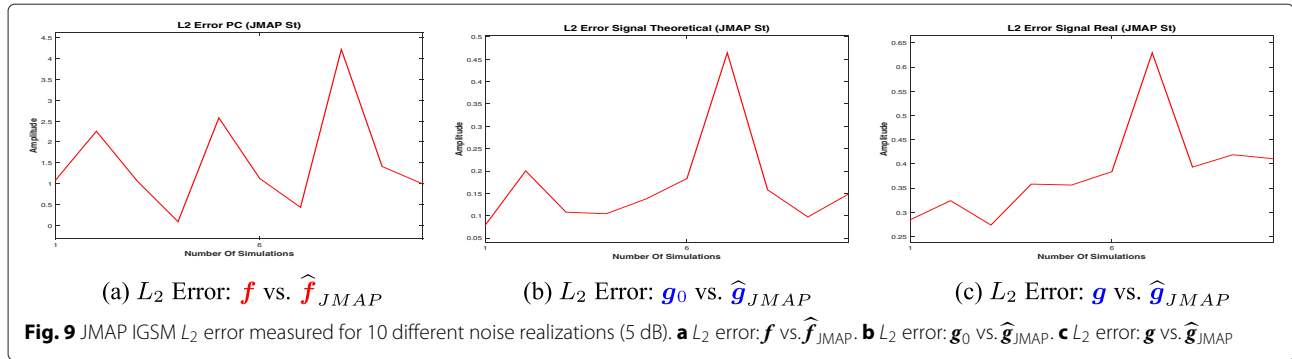
A comparison between the synthetic data and the PM (via VBA, partial separability) IGSM estimation is presented in Fig. 10. We compare the theoretical PC vector  $f$  with the PM (via VBA, partial separability) IGSM estimation  $\hat{f}_{PM}$



(Fig. 10a) and the corresponding reconstructed signal  $\hat{\mathbf{g}}_{PM}$  both with the theoretical signal  $\mathbf{g}_0$  (Fig. 10b) and the input signal  $\mathbf{g}$  (Fig. 10c).

In the case of the posterior mean estimation via VBA, both the PC estimation and theoretical signal  $\mathbf{g}_0$

reconstruction are very accurate (Fig. 10a, b). For the reconstruction of the theoretical signal  $\mathbf{g}_0$ , the  $L_2$  error norm is  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{PM}\|_2}{\|\mathbf{g}_0\|_2} = 0.0275$ . For the PC vector, the reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{PM}\|_2}{\|\mathbf{f}\|_2} = 0.0283$ . The



algorithm is converging to a sparse solution where all the non-zero peaks are detected. The residual error computed between  $\mathbf{g}$  and the reconstructed signal is consistent with the error considered in the model, 5 dB (Fig. 10c). During the algorithm, both inverse gamma shape parameters are constant (Eqs. (24b) and (24c)).

We present the convergence of the scale parameters  $\beta_\epsilon$  and  $\beta_f$ , (Fig. 11b, d), the convergence of  $\Sigma$  covariance matrix diagonal (Fig. 11c), and the convergence of the algorithm’s solution  $\mathbf{f}$ . For a better visualization of the PC convergence,  $\mathbf{f}$  is plotted as a vector (Fig. 11a). The color scale corresponding to each figure represents the iterations, showing a very fast convergence both for the parameters and hyperparameters involved in the model. All the estimations of the parameters and hyperparameters are superposed after the first ten iterations. In the previous paragraph, we have showed that the JMAP estimation for the proposed model is presenting high variations in terms of the error estimation and reconstruction. We show that for the PM estimation, the error variation is very small.

Figure 12a presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realization. The figure presents a very small variation of  $L_2$  PC vector error reconstruction, between 0.02215 and 0.0621. Very small variations corresponding to the  $L_2$  error reconstruction for the theoretical signal  $\mathbf{g}_0$  and signal  $\mathbf{g}$  are presented in Fig. 10b, c.

#### 5.1.4 PM (via VBA, full separability) IGSM 05 dB

The estimations for the full separability case are also accurate (Fig. 13).

Numerically, for the reconstruction of the theoretical signal  $\mathbf{g}_0$ , the  $L_2$  error norm is  $\delta\mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{PM}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.0247$ . For the PC vector, the reconstruction error is  $\delta\mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{PM}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.0234$ .

Figure 14a presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realization. The figure presents a very small variation of  $L_2$  PC vector error reconstruction, between 0.02 and 0.067. Very small

variations corresponding to the  $L_2$  error reconstruction for the theoretical signal  $\mathbf{g}_0$  and signal  $\mathbf{g}$  are presented in Fig. 14b, c.

#### 5.1.5 Methods comparison 05 dB

A comparison between the estimations corresponding to the IGSM proposed model is presented in Fig. 15c (JMAP estimator), d (PM via VBA, partial separability estimator), and e (PM via VBA, full separability estimator). As mentioned in Section 3, during this article, we adopted a Bayesian approach. However, other approaches are possible, via regularization. For this reason, we include a comparison with the Gaussian case (i.e., Gaussian prior), via the two estimators discussed, Fig. 15a (Gaussian model, JMAP estimator) and Fig. 15b (PM via VBA estimator). A comparison with the FFT is presented in Fig. 15f.

The  $L_2$  estimation error for the PC vector is very high for the two Gaussian models. Also, the estimations are not sparse. For the IGSM models, the JMAP estimator is providing a good estimation, but it is unstable. PM via VBA estimation, both partial and fully separable, provides very accurate stable estimations.

#### 5.1.6 Error comparison 05 dB

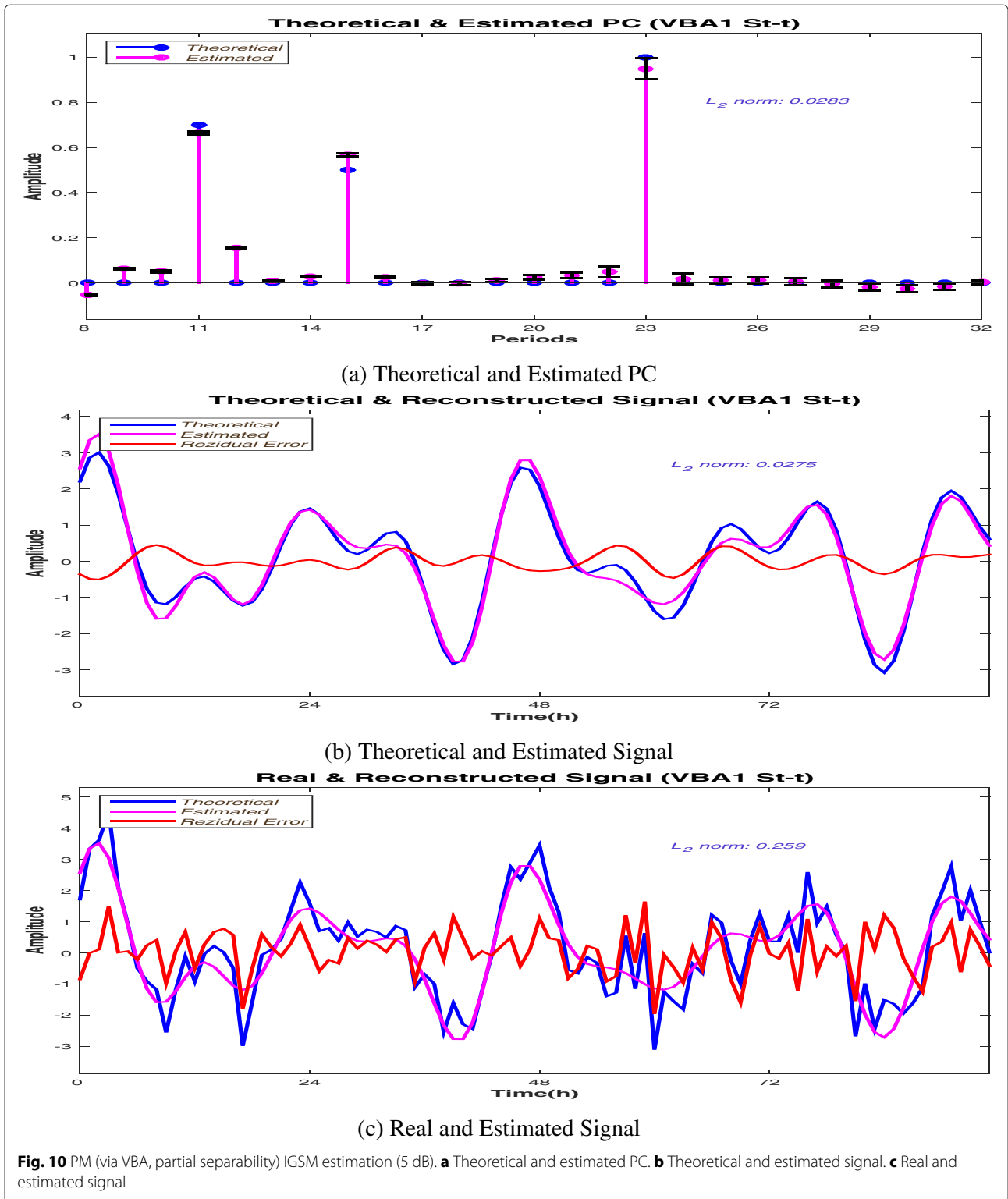
The  $L_2$  error measurement corresponding to the PC estimation, theoretical signal estimation, and signal estimation, for 10 different noise realization, is presented in the Fig. 16.

The  $L_2$  error corresponds to the PM via the VBA IGSM model corresponding to the PC vector estimation; Fig. 16a shows the performances of the proposed algorithm compared to the Gaussian model and the JMAP estimation for IGSM model.

## 5.2 Real data

This subsection is dedicated to the results corresponding to the real data, obtained in the experiments in chronobiology for cancer treatment. The particular experiment presented is realized on mice, investigating the tumor clock gene expression and the locomotor activity (rest-activity patterns) of KI/KI Per2::luc mouse, aged 10

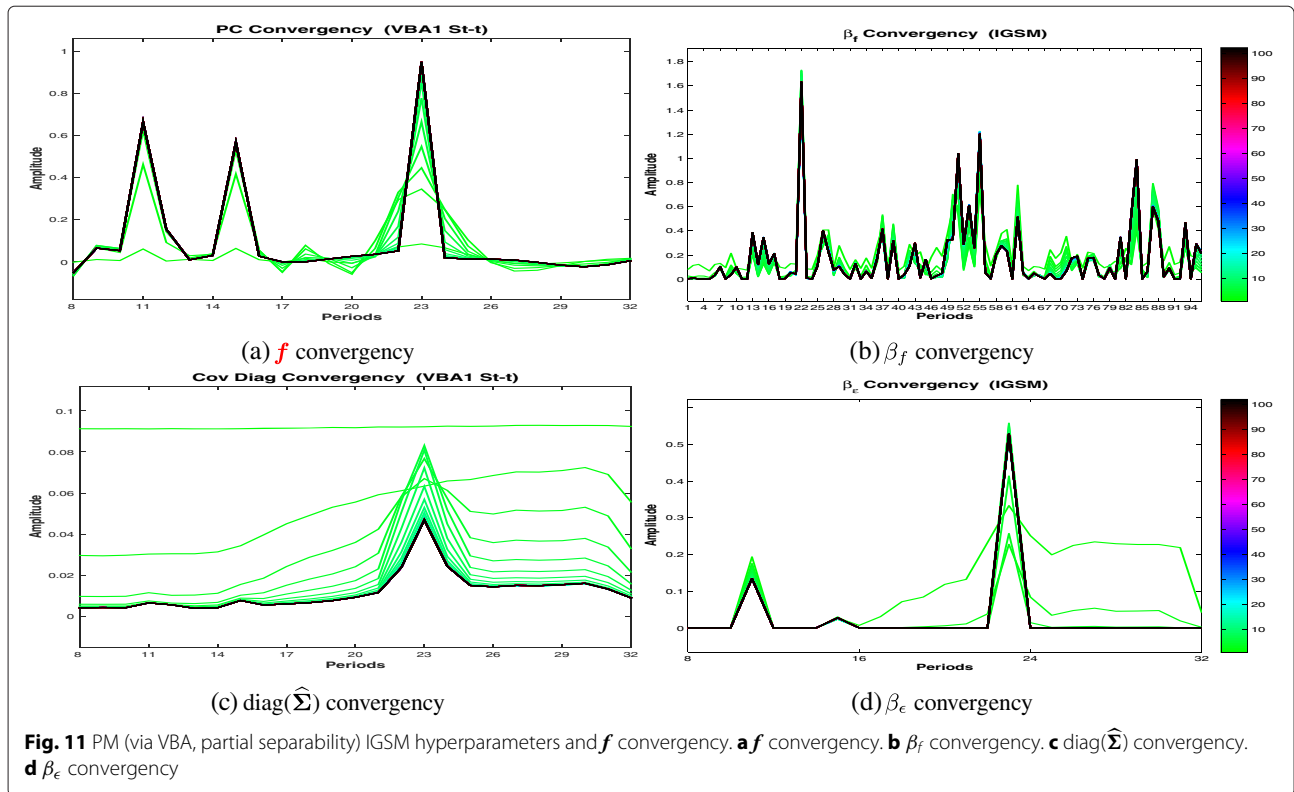




**Fig. 10** PM (via VBA, partial separability) IGSM estimation (5 dB). **a** Theoretical and estimated PC. **b** Theoretical and estimated signal. **c** Real and estimated signal

weeks, singly housed in RT-BIO and synchronized with LD 12:12 (i.e., 12 h of light, followed by 12 h of darkness). The signal considered in this section is representing the locomotor activity of the mouse, which is known to

be rhythmic. After the LD part of the signal, the mouse is kept in total darkness (DD) for 3 days, corresponding to the before-treatment part of the signal and then D-luciferin is loaded in subcutaneous implanted Alzet pump

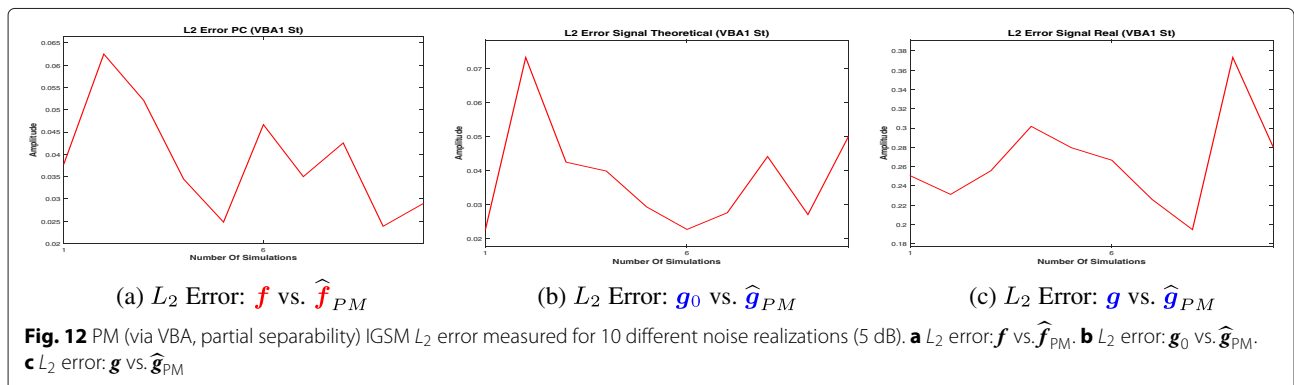


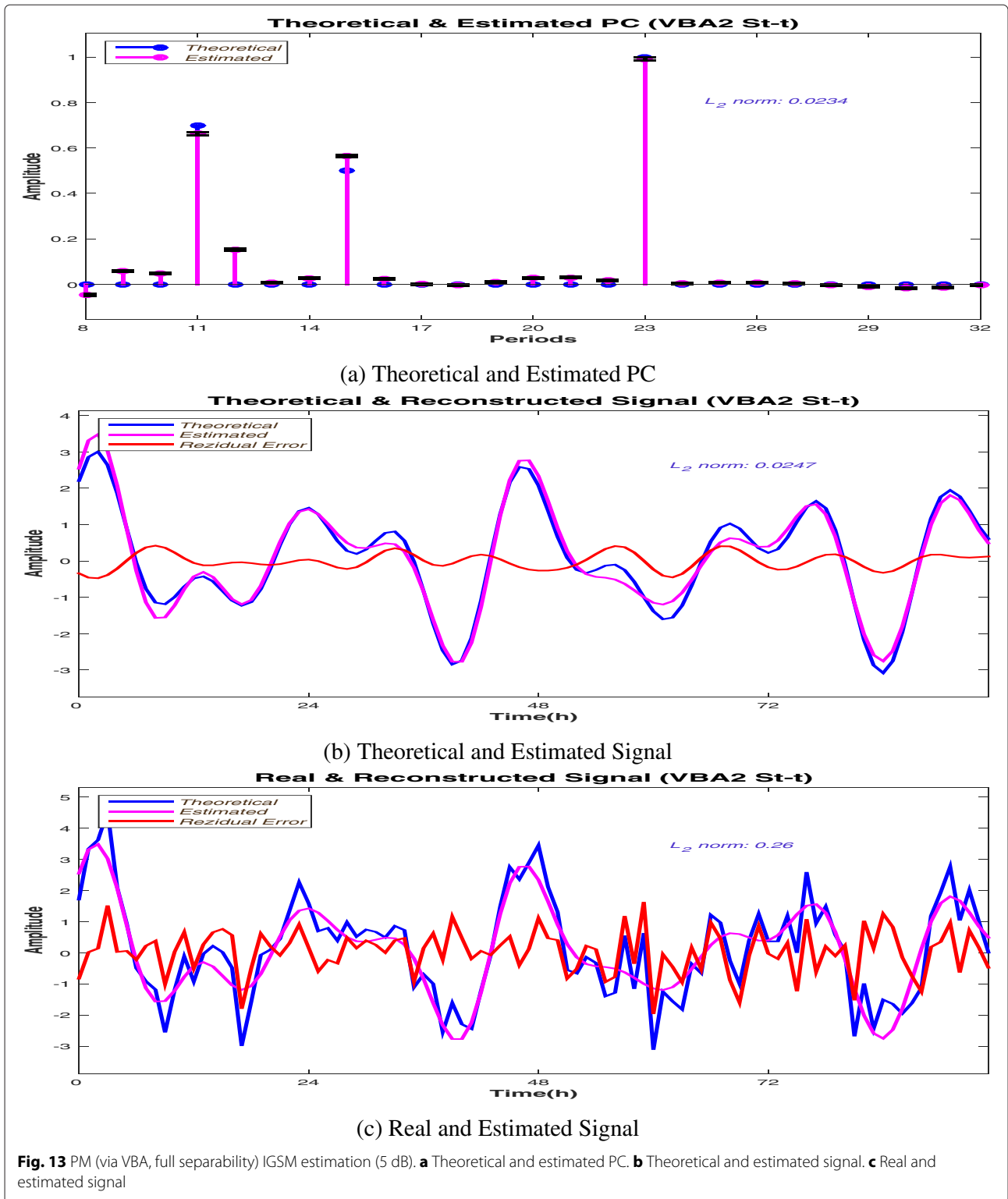
[90 mg/ml], recording for 5 days the signal corresponding to the during-treatment part of the signal. The last 2 days represent the after-treatment part of the signal. During the DD segment, the locomotor activity might be perturbed, due to the absence of the light-day regime and due to the treatment effects. Fig. 17a presents the raw data corresponding to the locomotor activity signal. The four segments of interest are indicated in the figure. The raw data signal was sampled every minute. The stability of the period during all four segments is verified using the classical FFT method and the proposed algorithm, PM via VBA, partial separability. For the segments corresponding to the LD and during treatment, we have considered the moving window strategy, i.e., we have considered

4-day-length signals shifted every day in order to verify the stability or the variability of the dominant period. For the four segments studied, we consider the mean-zero signals, normalized between  $[-10 : 10]$  and sampled every hour: the LD segment (Fig. 17b), the DD-before segment (Fig. 17c), the DD-during segment (Fig. 17d), and the DD-after segment (Fig. 17e).

For the LD segment, 7 days are available. We compute the PC corresponding to the signal using the proposed method and also using the FFT.

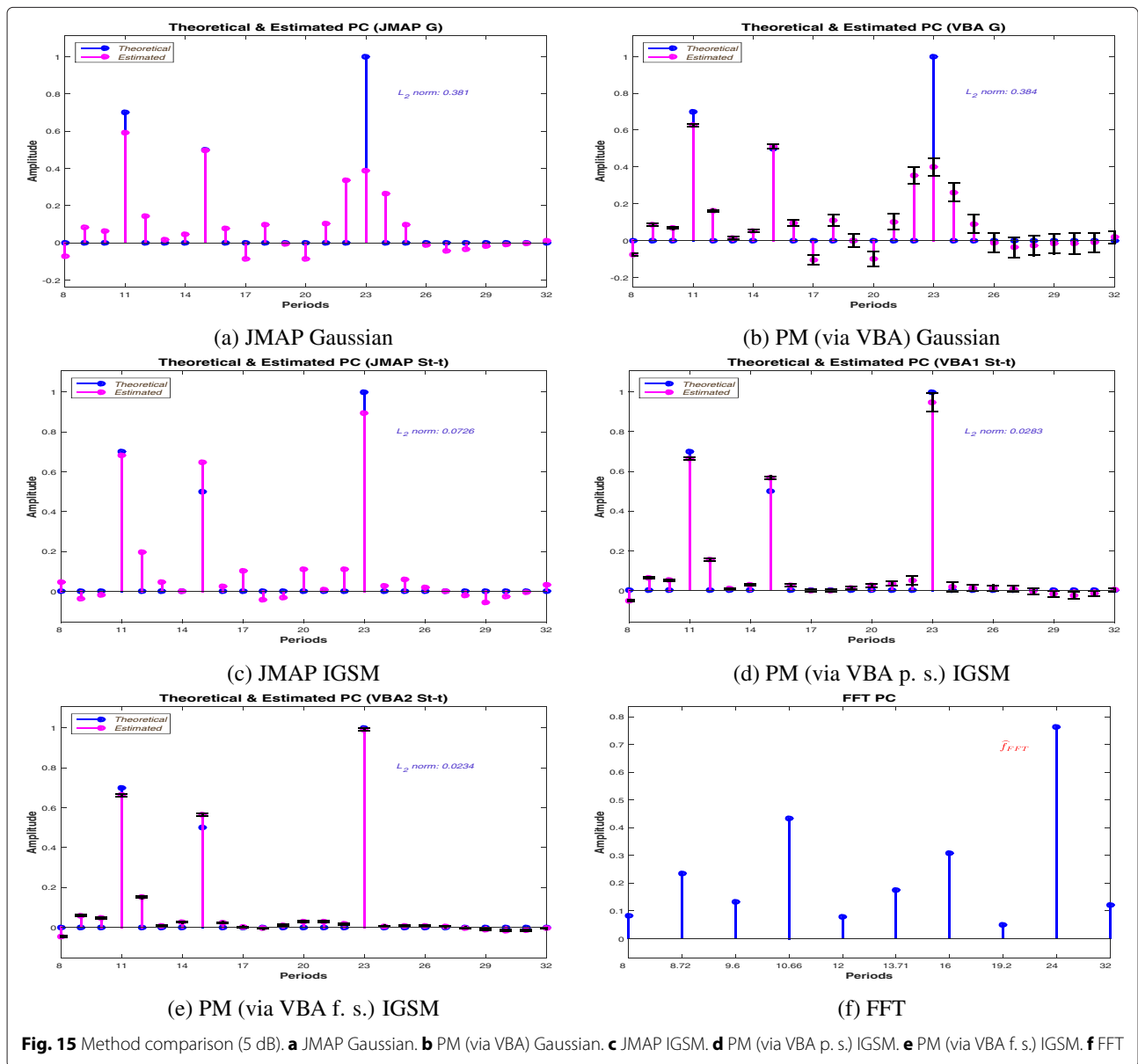
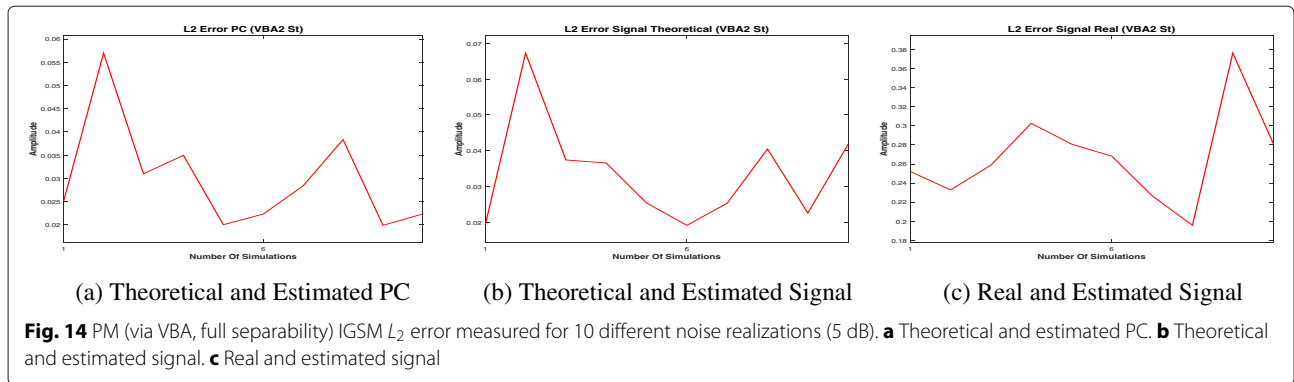
Via the FFT method, the dominant period is estimated at 24 h (Fig. 18c). Evidently, beside the incertitude associated with the FFT-estimated PC vector, the existence of other rhythms cannot be established, being difficult to

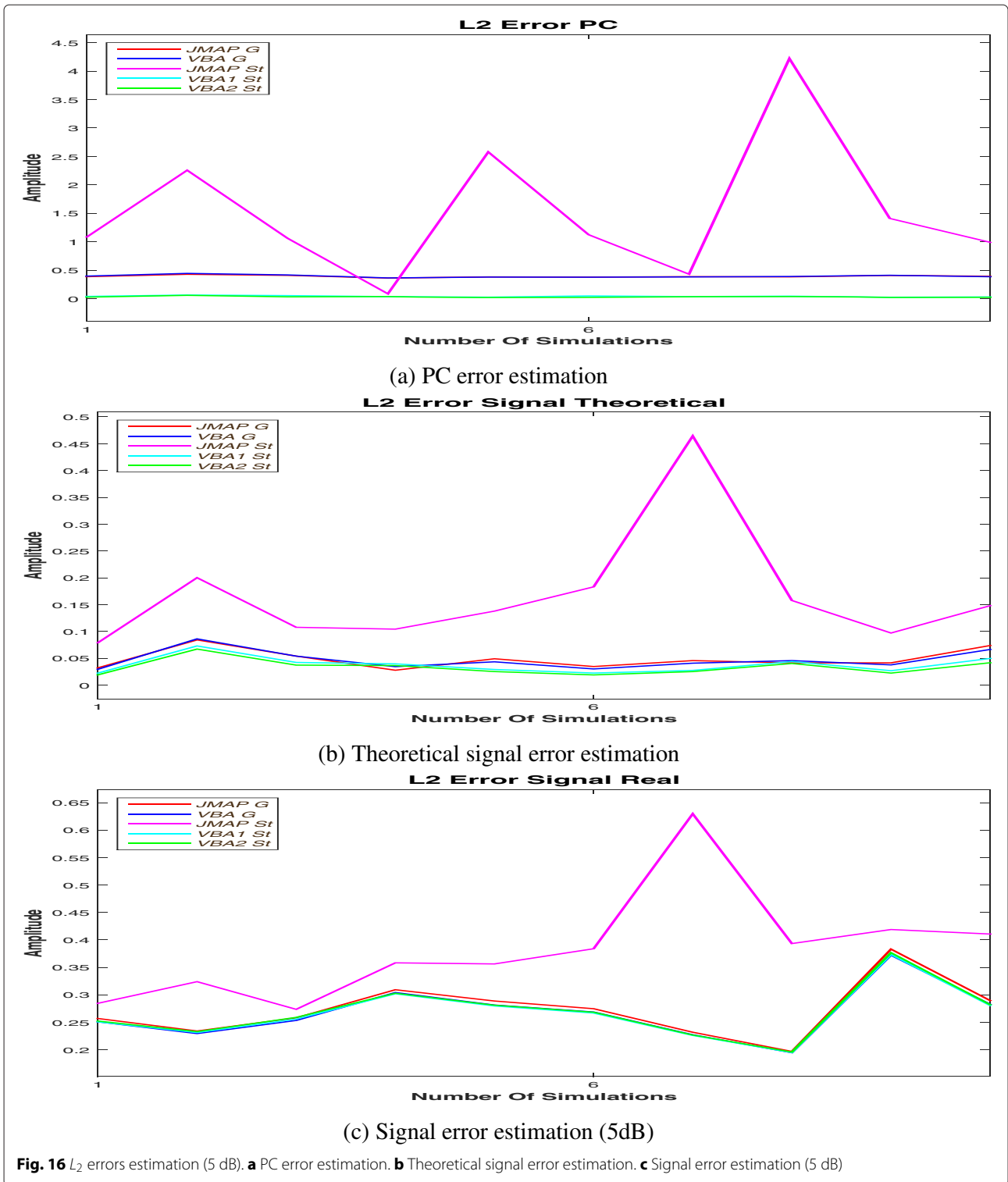




interpret all the peaks that appear PC vector. Via the proposed method, the estimated PC vector is a sparse one and the dominant period is estimated at 23 h (Fig. 18b). We note that via the proposed method, there is no uncertainty

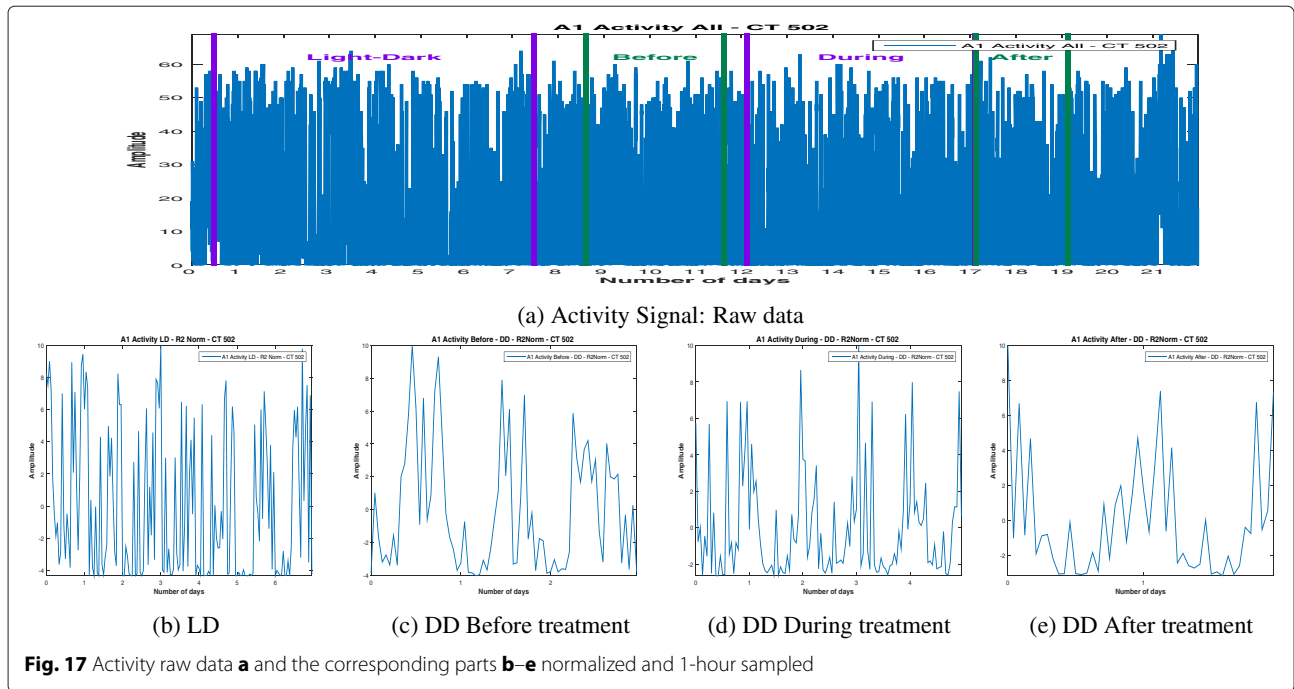
concerning the biological phenomena. We consider 4-day-length signals (windows) from the available signal, with a shift of 1 day and compute the PC via FFT and the proposed method.





The four windows considered are presented in Fig. 19a, d, g, j, and all four windows present a 24-h dominant period, via the FFT estimation (Fig. 19c, f, i, l). Via the proposed method, we obtain sparse PC vectors, showing a variability of the dominant period, between 23 and

24 h. A comparison between the proposed method and the FFT method is presented in Fig. 20, showing the stability of the dominant period established by the FFT method (Fig. 20b) and the variability established by the proposed method (Fig. 20a) (the  $x$ -axis represents the periods



inside the circadian domain and the  $y$ -axis represents the windows).

For the DD period, we consider first the before-treatment segment. A 3-day-length signal is available. The estimate PC vector, corresponding to the proposed method and FFT method is presented in Fig. 21.

Via the FFT, the highest peak is set at 24 h and the next highest peak is set at 8 h and the next highest peak is set at 8 h. Given the short length of the signal, 3 days, and the limitations of the FFT method, all the values inside the interval (18, 36) except 24 are not present in the estimated vector; so the values are uncertain. Via the proposed method, the dominant period is set at 22 h. For the during-treatment part of the data, a 5-day-length signal is available.

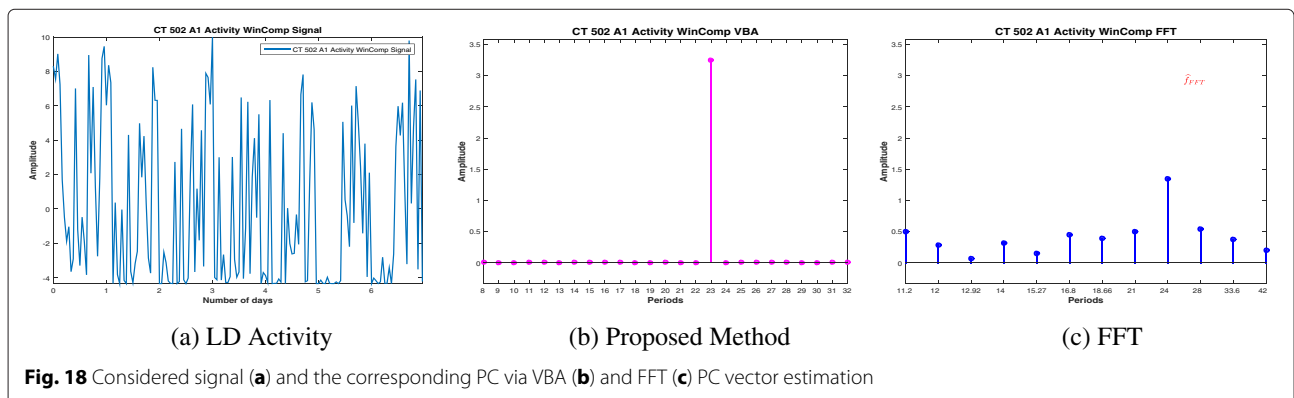
Via the proposed method, the estimated PC vector is a sparse vector, in accordance with the model, and the dominant period is estimated at 25 h (Fig. 22b). For the

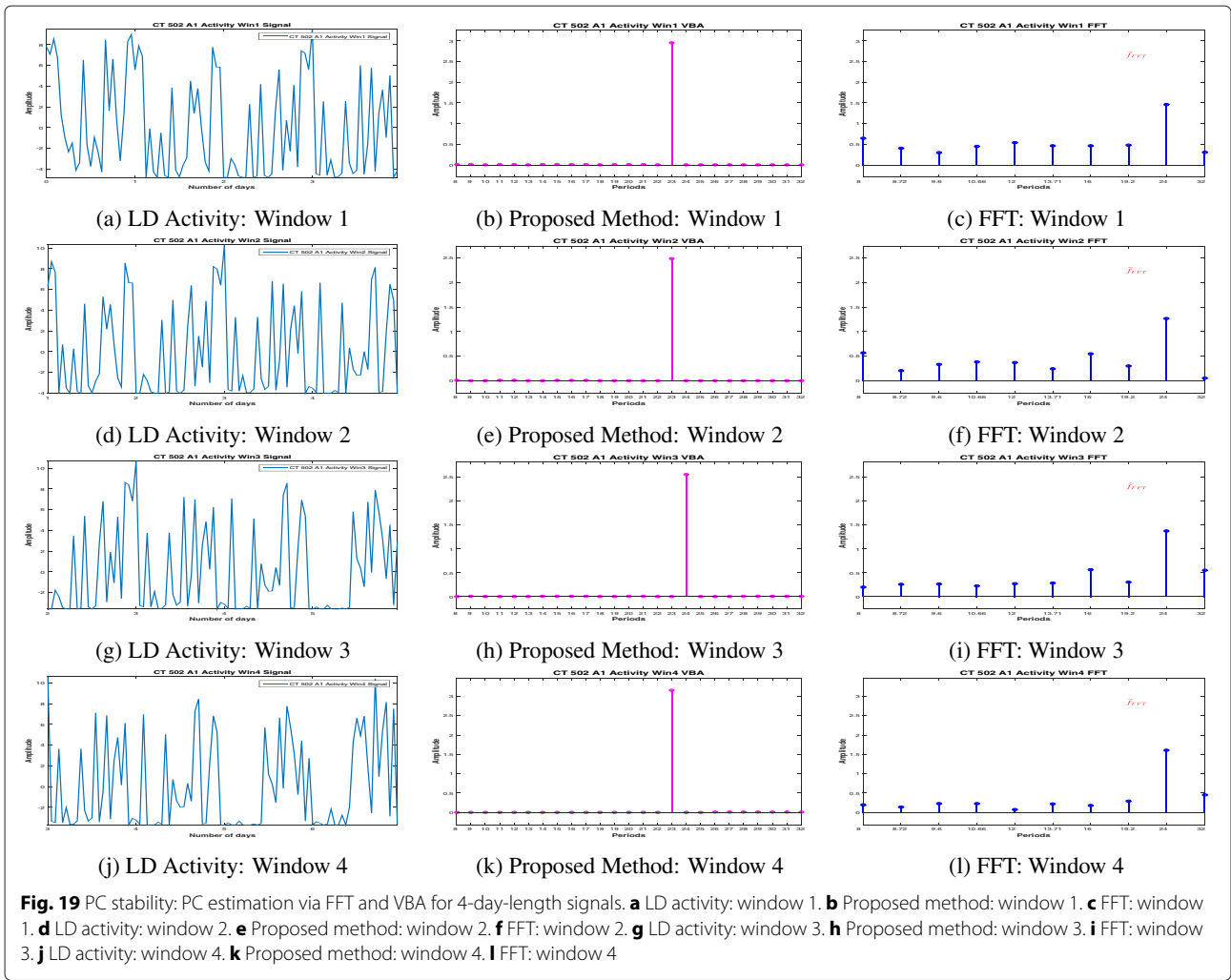
FFT-estimated PC vector, the dominant period is set at 24 h (Fig. 22c). Considering 4-day-length signals, we analyse the stability of the dominant period. Figures 23 and 24 show a stability of the dominant period established by the FFT method and a variability of the dominant period established by the proposed method.

For the after treatment part, only a 2-day-length signal is available. The FFT method is establishing a 24-h dominant period (Fig. 25c) while via the proposed method, the PC vector contains only one period, at 25 h (Fig. 25b).

### 6 Conclusions

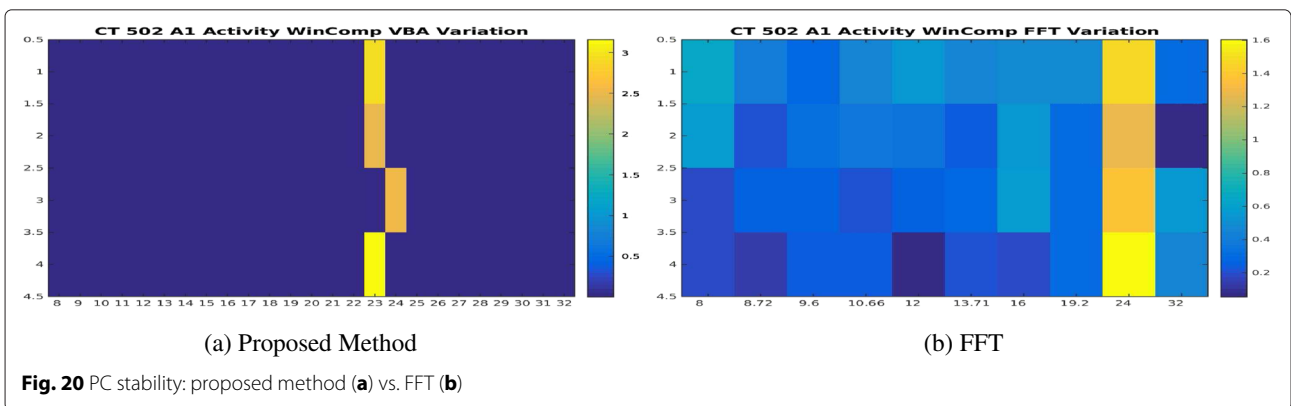
In this article, we have proposed a new method for a precise estimation of the PC vector for biomedical signals, based on the general Bayesian inference and using a hierarchical model with sparsity enforcing prior. The prior considered was a Student's  $t$  distribution expressed as the

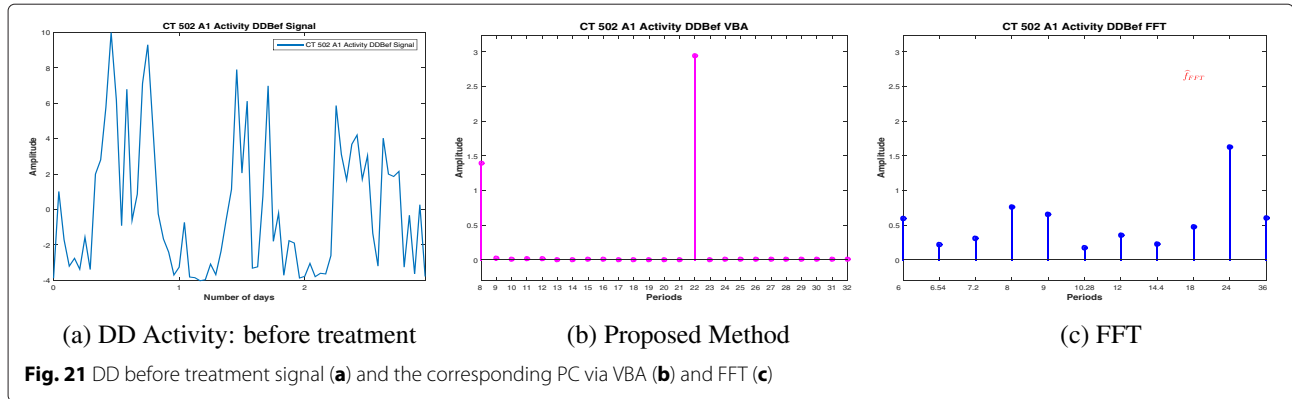




marginal of an infinite Gaussian scale mixture. The context of our work were the short signals relative to the prior knowledge for the dominant period (4-day signals and 24-h period). In Subsection 5.2, we applied the proposed method also for 2- and 3-day-length signals. The objective was to develop a method that can improve the

precision given by the FFT method and also to account for the possible effects of the measurement errors and the uncertainties. The method was tested first on synthetic data, in order to be validated. The algorithms corresponding to the Gaussian model (JMAP and PM estimators) fail to accurately reconstruct the sparse theoretical PC vector.





When using the JMAP estimator for the IGSM hierarchical model, the estimation is unstable. The error vectors corresponding to the JMAP-IGSM estimation (Fig. 16a, b) are showing the drawbacks of the method. Both PM-IGSM models accurately estimate the theoretical PC vector, (Fig. 10a, SNR = 05 dB). The comparison between the reconstructed signal and the theoretical input (Fig. 10b, SNR = 05 dB) and the comparison between the reconstructed signal and the noised input (Fig. 10c, SNR = 05 dB) show a good reconstruction and a good residual error, consistent with the considered added noise for the noised signal  $\mathbf{g}$ . These algorithms allow the estimation of the covariance matrix. The convergence of  $\mathbf{f}$  and hyperparameters is showing a fast convergence of the PM algorithms. The proposed method, PM via VBA, IGSM model was validated for a different set of data, at different ratios of noise, and the estimate was accurate in all the cases. For the real data, a comparison between the outputs is impossible. We have presented a comparison between the PC estimate corresponding to the PM-IGSM algorithm and the FFT estimate. The proposed method offers more precision compared to the FFT and is able to select

the peaks corresponding to the biological phenomena. Via the proposed method, the conclusion imposed by the FFT method that the considered experiment presents a stability of the dominant period at 24 h is invalidated, showing a variation of the dominant period between 22 and 25 h.

**Appendices**

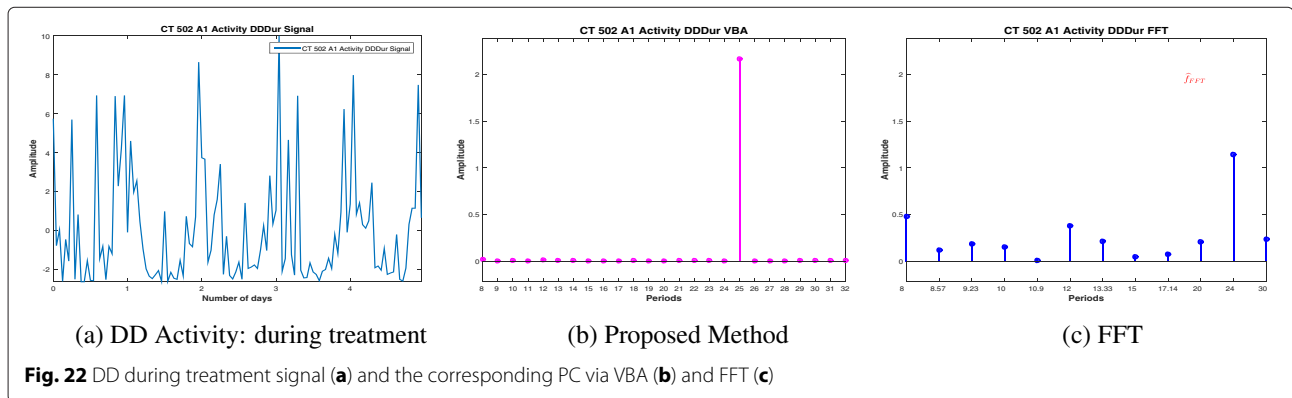
**Appendix 1**

**Computations for JMAP estimation**

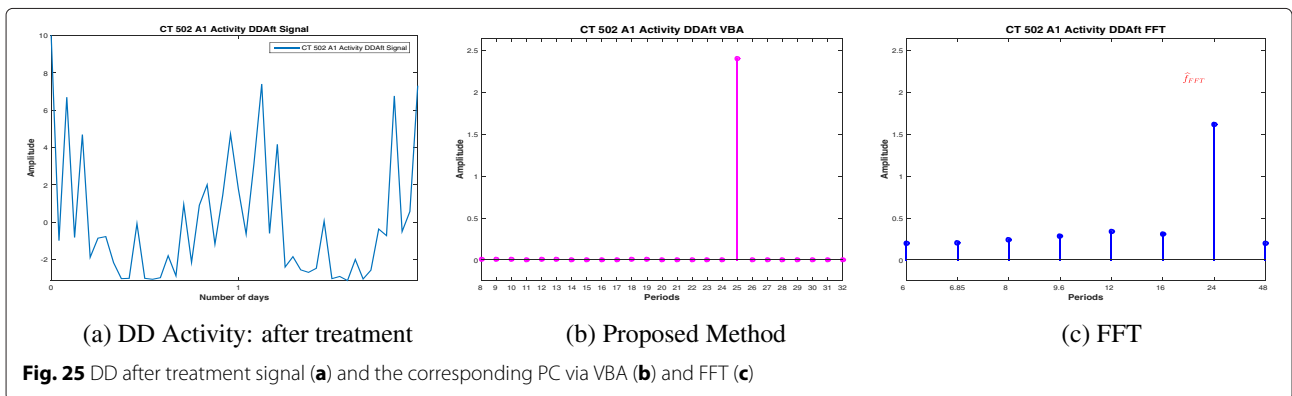
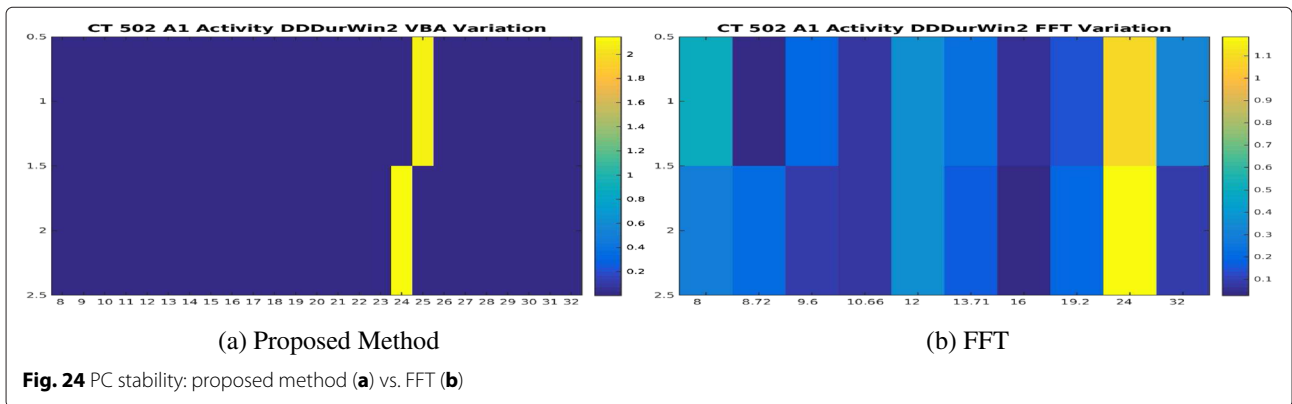
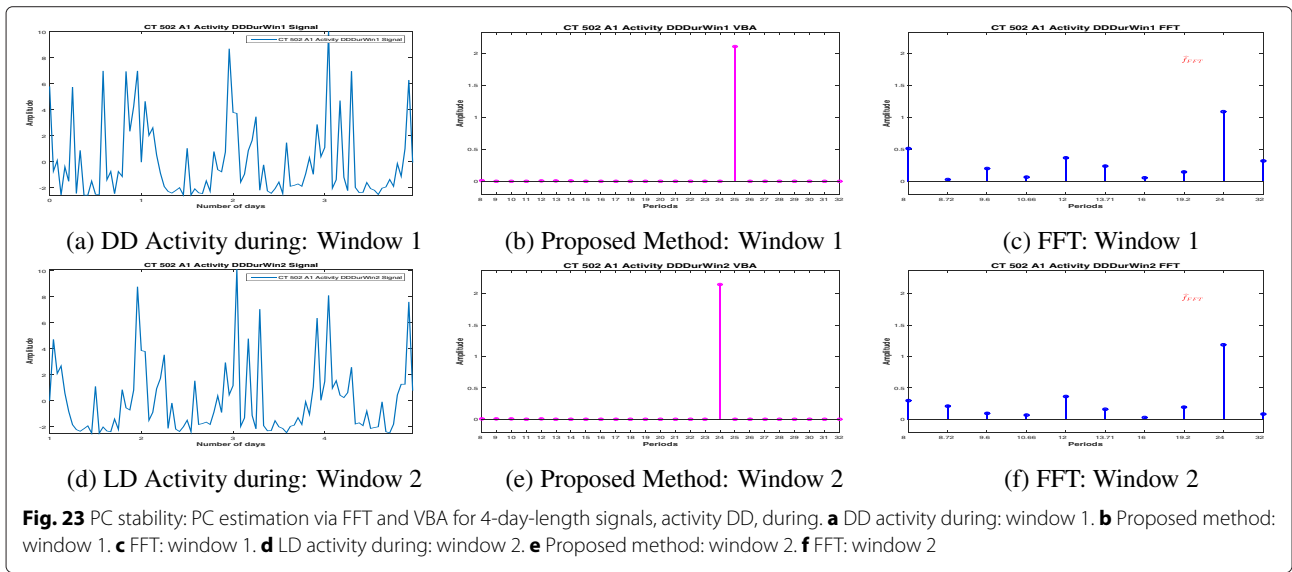
This section presents the computation for the joint MAP estimation (Subsection 4.1). The estimation is done via alternate optimization. The criterion is  $\mathcal{L}(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f) = -\ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$ , and  $p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  is defined in Eq. (15).

• With respect to  $\mathbf{f}$ :

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{f}, \hat{\mathbf{v}}_\epsilon, \hat{\mathbf{v}}_f)}{\partial \mathbf{f}} = 0 &\Leftrightarrow \frac{\partial}{\partial \mathbf{f}} \left( \|\mathbf{V}_\epsilon^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 + \|\mathbf{V}_f^{-\frac{1}{2}}\mathbf{f}\|^2 \right) = 0 \\ &\Leftrightarrow -\mathbf{H}^T \mathbf{V}_\epsilon^{-1}(\mathbf{g} - \mathbf{H}\mathbf{f}) + \mathbf{V}_f^{-1}\mathbf{f} = 0 \\ &\Leftrightarrow \left[ \mathbf{H}^T \mathbf{V}_\epsilon^{-1} \mathbf{H} + \mathbf{V}_f^{-1} \right] \mathbf{f} = \mathbf{H}^T \mathbf{V}_\epsilon^{-1} \mathbf{g} \\ &\Rightarrow \hat{\mathbf{f}}_{\text{JMAP}} = \left[ \mathbf{H}^T \mathbf{V}_\epsilon^{-1} \mathbf{H} + \mathbf{V}_f^{-1} \right]^{-1} \mathbf{H}^T \mathbf{V}_\epsilon^{-1} \mathbf{g} \end{aligned}$$







- With respect to  $v_{\epsilon_i}, i \in \{1, 2, \dots, N\}$ :

$$\begin{aligned} \frac{\partial \mathcal{L}(\hat{\mathbf{f}}, \mathbf{v}_{\epsilon}, \hat{\mathbf{v}}_f)}{\partial v_{\epsilon_i}} = 0 &\Leftrightarrow \frac{\partial}{\partial v_{\epsilon_i}} \left( \frac{1}{2} \ln \det(\mathbf{V}_{\epsilon}) + \frac{1}{2} \|\mathbf{V}_{\epsilon}^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 + (\alpha_{\epsilon 0} + 1) \ln v_{\epsilon_i} + \beta_{\epsilon 0} v_{\epsilon_i}^{-1} \right) = 0 \\ &\Leftrightarrow \frac{\partial}{\partial v_{\epsilon_i}} \left( \left( \alpha_{\epsilon 0} + 1 + \frac{1}{2} \right) \ln v_{\epsilon_i} + \left[ \beta_{\epsilon 0} + \frac{1}{2} (g_i - \mathbf{H}_i \mathbf{f})^2 \right] v_{\epsilon_i}^{-1} \right) = 0 \\ &\Leftrightarrow \left( \alpha_{\epsilon 0} + 1 + \frac{1}{2} \right) v_{\epsilon_i} - \left( \beta_{\epsilon 0} + \frac{1}{2} (g_i - \mathbf{H}_i \mathbf{f})^2 \right) = 0 \\ &\Rightarrow \hat{v}_{\epsilon_i \text{IMAP}} = \frac{\beta_{\epsilon 0} + \frac{1}{2} (g_i - \mathbf{H}_i \mathbf{f})^2}{\alpha_{\epsilon 0} + 1 + \frac{1}{2}} \end{aligned}$$

- With respect to  $v_f, j \in \{1, 2, \dots, M\}$ :

$$\begin{aligned} \frac{\partial \mathcal{L}(\hat{\mathbf{f}}, \hat{\mathbf{v}}_{\epsilon}, \mathbf{v}_f)}{\partial v_{f_j}} = 0 &\Leftrightarrow \frac{\partial}{\partial v_{f_j}} \left( \frac{1}{2} \ln \det(\mathbf{V}_f) + \frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}\|^2 + (\alpha_{f 0} + 1) \ln v_{f_j} + \beta_{f 0} v_{f_j}^{-1} \right) = 0 \\ &\Leftrightarrow \frac{\partial}{\partial v_{f_j}} \left( \left[ \alpha_{f 0} + 1 + \frac{1}{2} \right] \ln v_{f_j} + \left[ \beta_{f 0} + \frac{f_j^2}{2} \right] v_{f_j}^{-1} \right) = 0 \\ &\Leftrightarrow \left( \alpha_{f 0} + 1 + \frac{1}{2} \right) v_{f_j} - \left( \beta_{f 0} + \frac{f_j^2}{2} \right) = 0 \\ &\Rightarrow \hat{v}_{f_j \text{IMAP}} = \frac{\beta_{f 0} + \frac{f_j^2}{2}}{\alpha_{f 0} + 1 + \frac{1}{2}} \end{aligned}$$

## Appendix 2

### Computations for PM estimation via VBA, partial separability

This section presents the computation for the PM estimation, via VBA, partial separability (Subsection 4.2). The analytical expression of the logarithm is as follows:

$$\begin{aligned} \ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_f | \mathbf{g}) = & -\frac{1}{2} \ln \det(\mathbf{V}_{\epsilon}) - \frac{1}{2} \|\mathbf{V}_{\epsilon}^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 - \frac{1}{2} \ln \det(\mathbf{V}_f) - \frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}\|^2 \\ & - \sum_{i=1}^N (\alpha_{\epsilon 0} + 1) \ln v_{\epsilon_i} - \sum_{i=1}^N \beta_{\epsilon 0} v_{\epsilon_i}^{-1} - \sum_{j=1}^M (\alpha_{f 0} + 1) \ln v_{f_j} - \sum_{j=1}^M \beta_{f 0} v_{f_j}^{-1} + C \end{aligned} \quad (33)$$

- **Expression of  $q_1(\mathbf{f})$ :**

The proportionality relation concerning  $q_1(\mathbf{f})$  established in Eq. (22) refers to  $\mathbf{f}$ , so in the expression of  $\ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_f | \mathbf{g})$ , all the terms free of  $\mathbf{f}$  can be regarded as constants:

$$\langle \ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_f | \mathbf{g}) \rangle_{q_2(\mathbf{v}_{\epsilon}) q_3(\mathbf{v}_f)} = \left\langle C - \frac{1}{2} \|\mathbf{V}_{\epsilon}^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 - \frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}\|^2 \right\rangle_{q_2(\mathbf{v}_{\epsilon}) q_3(\mathbf{v}_f)}$$

leading to:

$$\langle \ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_f | \mathbf{g}) \rangle_{q_2(\mathbf{v}_{\epsilon}) q_3(\mathbf{v}_f)} = C - \frac{1}{2} \left\langle \|\mathbf{V}_{\epsilon}^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 \right\rangle_{q_2(\mathbf{v}_{\epsilon})} - \frac{1}{2} \left\langle \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}\|^2 \right\rangle_{q_3(\mathbf{v}_f)} \quad (34)$$

Considering the notation introduced in (10) corresponding to  $\mathbf{V}_{\epsilon}$  and denoting the  $i$ th line of the matrix  $\mathbf{H}$  with  $\mathbf{H}_i$ ,  $i \in \{1, 2, \dots, N\}$ , we write:

$$\mathbf{V}_{\epsilon}^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f}) = [v_{\epsilon_1}^{-1/2}(g_1 - \mathbf{H}_1 \mathbf{f}) \dots v_{\epsilon_i}^{-1/2}(g_i - \mathbf{H}_i \mathbf{f}) \dots v_{\epsilon_N}^{-1/2}(g_N - \mathbf{H}_N \mathbf{f})]^T \quad (35)$$

so the norm is written as:

$$\|\mathbf{V}_{\epsilon}^{-\frac{1}{2}}(\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 = \sum_{i=1}^N v_{\epsilon_i}^{-1} (g_i - \mathbf{H}_i \mathbf{f})^2 \quad (36)$$

Introducing the notations:

$$\begin{aligned} \widetilde{v}_{\epsilon_i}^{-1} &= \int v_{\epsilon_i}^{-1} q_{2i}(v_{\epsilon_i}) dv_{\epsilon_i} ; \quad \widetilde{\mathbf{v}}_{\epsilon}^{-1} = \left[ \widetilde{v}_{\epsilon_1}^{-1} \dots \widetilde{v}_{\epsilon_i}^{-1} \dots \widetilde{v}_{\epsilon_N}^{-1} \right]^T ; \\ \widetilde{\mathbf{V}}_{\epsilon}^{-1} &= \text{diag} \left( \widetilde{\mathbf{v}}_{\epsilon}^{-1} \right) \end{aligned} \quad (37)$$

we can write:

$$\begin{aligned} \left\langle \left\| \mathbf{V}_{\epsilon}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 \right\rangle_{q_2(\mathbf{v}_{\epsilon})} &= \sum_{i=1}^N \widetilde{v}_{\epsilon_i}^{-1} (g_i - \mathbf{H}_i \mathbf{f})^2 \\ &= \left\| \left( \widetilde{\mathbf{V}}_{\epsilon}^{-1} \right)^{1/2} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 \end{aligned} \quad (38)$$

Introducing the notation

$$\begin{aligned} \widetilde{v}_{f_j}^{-1} &= \int v_{f_j}^{-1} q_{3j}(v_{f_j}) dv_{f_j} ; \quad \widetilde{\mathbf{v}}_{\mathbf{f}}^{-1} = \left[ \widetilde{v}_{f_1}^{-1} \dots \widetilde{v}_{f_j}^{-1} \dots \widetilde{v}_{f_M}^{-1} \right]^T ; \\ \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} &= \text{diag} \left( \widetilde{\mathbf{v}}_{\mathbf{f}}^{-1} \right) \end{aligned} \quad (39)$$

we can write:

$$\left\langle \left\| \mathbf{V}_{\mathbf{f}}^{-\frac{1}{2}} \mathbf{f} \right\|^2 \right\rangle_{q_3(\mathbf{v}_{\mathbf{f}})} = \left\| \left( \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right)^{\frac{1}{2}} \mathbf{f} \right\|^2 \quad (40)$$

Finally from (34), (38), and (40), for the expression of  $\langle \ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_{\mathbf{f}} | \mathbf{g}) \rangle_{q_2(\mathbf{v}_{\epsilon}) q_3(\mathbf{v}_{\mathbf{f}})}$ , we have:

$$\begin{aligned} \langle \ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_{\mathbf{f}} | \mathbf{g}) \rangle_{q_2(\mathbf{v}_{\epsilon}) q_3(\mathbf{v}_{\mathbf{f}})} &= C - \frac{1}{2} \left\| \left( \widetilde{\mathbf{V}}_{\epsilon}^{-1} \right)^{1/2} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 \\ &\quad - \frac{1}{2} \left\| \left( \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right)^{\frac{1}{2}} \mathbf{f} \right\|^2 \end{aligned} \quad (41)$$

and via the first proportionality from (22) and the notation:

$$J(\mathbf{f}) = \left\| \left( \widetilde{\mathbf{V}}_{\epsilon}^{-1} \right)^{1/2} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 + \left\| \left( \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right)^{\frac{1}{2}} \mathbf{f} \right\|^2 \quad (42)$$

the probability  $q_1(\mathbf{f})$  can be expressed by the following proportionality:

$$q_1(\mathbf{f}) \propto \left\{ -\frac{1}{2} J(\mathbf{f}) \right\} \quad (43)$$

The criterion  $J(\mathbf{f})$  introduced in Eq. (42) is quadratic in  $\mathbf{f}$ . Equation 43 establishes a proportionality relation between  $q_1(\mathbf{f})$  and an exponential function having as argument a quadratic criterion. This leads to the following:

**Intermediate conclusion 1.** The probability distribution function  $q_1(\mathbf{f})$  is a multivariate normal distribution.

Of course, the mean is given by the solution that minimizes the criterion  $J(\mathbf{f})$ , i.e., the solution of the equation  $\frac{\partial J(\mathbf{f})}{\partial \mathbf{f}} = 0$  (and in particular, this is the same criterion that

arrived in the MAP estimation technique for  $\mathbf{f}$ , with some formal differences):

$$\frac{\partial J(\mathbf{f})}{\partial \mathbf{f}} = 0 \Rightarrow \widehat{\mathbf{f}}_{\text{PM}} = \left( \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} + \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right)^{-1} \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} \quad (44)$$

The corresponding covariance matrix is computed by identification. On the one hand, we have the following relation:

$$\mathcal{N}(\mathbf{f} | \widehat{\mathbf{f}}_{\text{PM}}, \widehat{\Sigma}) \propto (\det(\widehat{\Sigma}))^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{f} - \widehat{\mathbf{f}}_{\text{PM}})^T \widehat{\Sigma}^{-1} (\mathbf{f} - \widehat{\mathbf{f}}_{\text{PM}}) \right\} \quad (45)$$

On the other hand, we have the following proportionality, given by Eq. (43):

$$\mathcal{N}(\mathbf{f} | \widehat{\mathbf{f}}_{\text{PM}}, \widehat{\Sigma}) \propto q_1(\mathbf{f}) \propto \exp \left\{ -\frac{1}{2} J(\mathbf{f}) \right\} \quad (46)$$

So, the covariance matrix  $\widehat{\Sigma}$  must respect the following relation:

$$\left( \mathbf{f} - \widehat{\mathbf{f}}_{\text{PM}} \right)^T \widehat{\Sigma}^{-1} \left( \mathbf{f} - \widehat{\mathbf{f}}_{\text{PM}} \right) \equiv J(\mathbf{f}), \quad (47)$$

where the sign  $\equiv$  represents an equality between the two terms until a free  $\mathbf{f}$  term. If we consider the covariance matrix

$$\widehat{\Sigma} = \left( \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} + \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right)^{-1} \quad (48)$$

we have the following equalities:

$$\begin{aligned} \left( \mathbf{f} - \widehat{\mathbf{f}}_{\text{PM}} \right)^T \widehat{\Sigma}^{-1} \left( \mathbf{f} - \widehat{\mathbf{f}}_{\text{PM}} \right) &= \left( \mathbf{f} - \widehat{\Sigma} \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} \right)^T \widehat{\Sigma}^{-1} \\ &\quad \left( \mathbf{f} - \widehat{\Sigma} \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} \right) \\ &= \left( \mathbf{f}^T - \mathbf{g}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} \widehat{\Sigma} \right) \\ &\quad \left( \widehat{\Sigma}^{-1} \mathbf{f} - \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} \right) \\ &= \mathbf{f}^T \left( \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} + \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right) \\ &\quad \mathbf{f} - 2 \mathbf{f}^T \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} + C, \end{aligned} \quad (49)$$

where we have used the equality  $\mathbf{f}^T \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} = \mathbf{g}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} \mathbf{f}$ , as a consequence of the fact that one term is the transpose of the other and the term is a scalar. We also used the fact that  $\widehat{\Sigma} = \widehat{\Sigma}^T$  and  $\mathbf{g}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} \left( \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} + \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right)^{-1} \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g}$  was viewed as a constant  $C$ . We also have the following equalities:

$$\begin{aligned} J(\mathbf{f}) &= \left\| \left( \widetilde{\mathbf{V}}_{\epsilon}^{-1} \right)^{1/2} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 + \left\| \left( \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right)^{\frac{1}{2}} \mathbf{f} \right\|^2 \\ &= \left( \mathbf{g}^T - \mathbf{f}^T \mathbf{H}^T \right) \widetilde{\mathbf{V}}_{\epsilon}^{-1} (\mathbf{g} - \mathbf{H}\mathbf{f}) + \mathbf{f}^T \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \mathbf{f} \\ &= \mathbf{f}^T \left( \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} + \widetilde{\mathbf{V}}_{\mathbf{f}}^{-1} \right) \mathbf{f} - 2 \mathbf{f}^T \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} + C. \end{aligned} \quad (50)$$

Equations 49 and (50) show that equality imposed in (47) is verified with the covariance matrix defined as in (48). So, for the normal distribution  $\mathcal{N}(\mathbf{f}|\widehat{\mathbf{f}}, \widehat{\Sigma})$  proportional to  $q_1(\mathbf{f})$ , we have the following parameters:

$$q_1(\mathbf{f}) = \mathcal{N}(\mathbf{f}|\widehat{\mathbf{f}}_{\text{PM}}, \widehat{\Sigma}), \begin{cases} \widehat{\mathbf{f}}_{\text{PM}} = (\mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} + \widetilde{\mathbf{V}}_f^{-1})^{-1} \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{g} \\ \widehat{\Sigma} = (\mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon}^{-1} \mathbf{H} + \widetilde{\mathbf{V}}_f^{-1})^{-1} \end{cases} \quad (51)$$

• **Expression of  $q_{2i}(v_{\epsilon_i})$  :**

The proportionality relation concerning  $q_{2i}(v_{\epsilon_i})$  established in Eq. (22) refers to  $v_{\epsilon_i}$ , so in the expression of  $\ln p(\mathbf{f}, v_{\epsilon}, v_f | \mathbf{g})$ , all the terms free of  $v_{\epsilon_i}$  can be regarded as constants:

$$\begin{aligned} & \langle \ln p(\mathbf{f}, v_{\epsilon}, v_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}) q_{2-i}(v_{\epsilon_i}) q_3(v_f)} \\ &= C - \frac{1}{2} \langle \ln \det(\mathbf{V}_{\epsilon}) \rangle_{q_{2-i}(v_{\epsilon_i})} \\ & \quad - (\alpha_{\epsilon 0} + 1) \ln v_{\epsilon_i} \\ & \quad - \frac{1}{2} \left\langle \|\mathbf{V}_{\epsilon}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 \right\rangle_{q_1(\mathbf{f}) q_{2-i}(v_{\epsilon_i})} - \beta_{\epsilon 0} v_{\epsilon_i}^{-1} \end{aligned} \quad (52)$$

For the first integral, it is trivial to verify:

$$\langle \ln \det(\mathbf{V}_{\epsilon}) \rangle_{q_{2-i}(v_{\epsilon_i})} = C + \ln v_{\epsilon_i} \quad (53)$$

For the second integral, we have the following development:

$$\left\langle \|\mathbf{V}_{\epsilon}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 \right\rangle_{q_1(\mathbf{f}) q_{2-i}(v_{\epsilon_i})} = \left\langle \|\widetilde{\mathbf{V}}_{\epsilon-i}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 \right\rangle_{q_1(\mathbf{f})} \quad (54)$$

where we have introduced the following notations:

$$\begin{aligned} \widetilde{\mathbf{v}}_{\epsilon-i}^{-1} &= \left[ \widetilde{v}_{\epsilon 1}^{-1} \dots \widetilde{v}_{\epsilon i-1}^{-1} v_{\epsilon_i}^{-1} \widetilde{v}_{\epsilon i+1}^{-1} \dots \widetilde{v}_{\epsilon N}^{-1} \right]^T ; \\ \widetilde{\mathbf{V}}_{\epsilon-i}^{-1} &= \text{diag}(\widetilde{\mathbf{v}}_{\epsilon-i}^{-1}) \end{aligned} \quad (55)$$

Again, using the fact that  $q_1(\mathbf{f})$  is a multivariate normal distribution, we have:

$$\begin{aligned} \left\langle \|\widetilde{\mathbf{V}}_{\epsilon-i}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 \right\rangle_{q_1(\mathbf{f})} &= \|\widetilde{\mathbf{V}}_{\epsilon-i}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\widehat{\mathbf{f}}_{\text{PM}})\|^2 \\ & \quad + \text{Tr} \left( \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon-i}^{-1} \mathbf{H} \widehat{\Sigma} \right) \end{aligned} \quad (56)$$

and considering as constants all terms free of  $v_{\epsilon_i}$ , we have:

$$\begin{aligned} \|\widetilde{\mathbf{V}}_{\epsilon-i}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\widehat{\mathbf{f}}_{\text{PM}})\|^2 &= C + v_{\epsilon_i}^{-1} (g_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 ; \\ \text{Tr} \left( \mathbf{H}^T \widetilde{\mathbf{V}}_{\epsilon-i}^{-1} \mathbf{H} \widehat{\Sigma} \right) &= C + v_{\epsilon_i}^{-1} \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T \end{aligned} \quad (57)$$

where  $\mathbf{H}_i$  is the line  $i$  of the matrix  $\mathbf{H}$ , so we can conclude:

$$\begin{aligned} & \left\langle \|\mathbf{V}_{\epsilon}^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}\mathbf{f})\|^2 \right\rangle_{q_1(\mathbf{f}) q_{2-i}(v_{\epsilon_i})} \\ &= C + \left[ \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T + (g_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 \right] v_{\epsilon_i}^{-1} \end{aligned} \quad (58)$$

From (52) via (53) and (58), we get:

$$\begin{aligned} & \langle \ln p(\mathbf{f}, \mathbf{z}, v_{\epsilon}, v_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}) q_{2-i}(v_{\epsilon_i}) q_3(v_f)} \\ &= C - \left( \alpha_{\epsilon 0} + 1 + \frac{1}{2} \right) \ln v_{\epsilon_i} \\ & \quad \left( \beta_{\epsilon 0} + \frac{1}{2} \left[ \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T + (g_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 \right] \right) v_{\epsilon_i}^{-1} \end{aligned}$$

from which we can establish the proportionality corresponding to  $q_{2i}(v_{\epsilon_i})$ :

$$q_{2i}(v_{\epsilon_i}) \propto v_{\epsilon_i}^{-(\alpha_{\epsilon 0} + 1 + \frac{1}{2})} \exp \left\{ - \left( \beta_{\epsilon 0} + \frac{1}{2} \left[ \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T + (g_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 \right] \right) v_{\epsilon_i}^{-1} \right\} \quad (59)$$

Equation (59) leads to the following.

**Intermediate conclusion 2.** The probability distribution function  $q_{3i}(v_{\epsilon_i})$  is an inverse gamma distribution, with the parameters  $\alpha_{\epsilon_i}$  and  $\beta_{\epsilon_i}$ :

We can write:

$$q_{2i}(v_{\epsilon_i}) = \mathcal{IG}(v_{\epsilon_i} | \alpha_{\epsilon_i}, \beta_{\epsilon_i}), \begin{cases} \alpha_{\epsilon_i} = \alpha_{\epsilon 0} + \frac{1}{2} \\ \beta_{\epsilon_i} = \beta_{\epsilon 0} + \frac{1}{2} \left[ \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T + (g_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 \right] \end{cases} \quad (60)$$

• **Expression of  $q_{3j}(v_{f_j})$  :**

The proportionality relation concerning  $q_{3j}(v_{f_j})$  established in Eq. (22) refers to  $v_{f_j}$ , so in the expression of  $\ln p(\mathbf{f}, \mathbf{z}, v_{\epsilon}, v_f | \mathbf{g})$ , all the terms free of  $v_{f_j}$  can be regarded as constants:

$$\begin{aligned} & \langle \ln p(\mathbf{f}, v_{\epsilon}, v_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}) q_2(v_{\epsilon}) q_{3-j}(v_{f_j})} \\ &= -\frac{1}{2} \langle \ln \det(\mathbf{V}_f) \rangle_{q_{3-j}(v_{f_j})} - (\alpha_{f 0} + 1) \ln v_{f_j} \\ & \quad - \frac{1}{2} \left\langle \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}\|^2 \right\rangle_{q_1(\mathbf{f}) q_{3-j}(v_{f_j})} - \beta_{f 0} v_{f_j}^{-1} \end{aligned} \quad (61)$$

Considering all  $v_{f_j}$  free terms as constants, it is easy to verify:

$$\langle \ln \det(\mathbf{V}_f) \rangle_{q_{3-j}(v_{f_j})} = C + \ln v_{f_j} \quad (62)$$

For the second integral:

$$\left\langle \left\| (V_f)^{-\frac{1}{2}} f \right\|^2 \right\rangle_{q_1(f) q_{3-j}(v_j)} = \left\langle \left\| \left( \widetilde{V}_{f-i}^{-1} \right)^{\frac{1}{2}} f \right\|^2 \right\rangle_{q_1(f)} \tag{63}$$

where we have introduced the notations:

$$\begin{aligned} \widetilde{v}_{f-i}^{-1} &= \left[ \widetilde{v}_{f_1}^{-1} \dots \widetilde{v}_{f_{i-1}}^{-1} \widetilde{v}_{f_i}^{-1} \widetilde{v}_{f_{i+1}}^{-1} \dots \widetilde{v}_{f_N}^{-1} \right]^T ; \\ \widetilde{V}_{f-i}^{-1} &= \text{diag} \left( \widetilde{v}_{f-i}^{-1} \right) \end{aligned} \tag{64}$$

Considering the fact that  $q_1(f)$  was established as a multivariate normal distribution, we have:

$$\begin{aligned} \left\langle \left\| \left( \widetilde{V}_{f-i}^{-1} \right)^{\frac{1}{2}} f \right\|^2 \right\rangle_{q_1(f)} &= \left\| \left( \widetilde{V}_{f-i}^{-1} \right)^{\frac{1}{2}} \widehat{f}_{PM} \right\|^2 + \text{Tr} \left( \widetilde{V}_{f-i}^{-1} \widehat{\Sigma} \right) \\ &= C + v_{f_i}^{-1} \left( \widehat{f}_{PM}^2 + \widehat{\Sigma}_{jj} \right) \end{aligned} \tag{65}$$

From (61) via (62) and (65), we get:

$$\begin{aligned} \langle \ln p(f, v_\epsilon, v_f | g) \rangle_{q_1(f) q_2(v_\epsilon) q_{3-j}(v_j)} &= - \left( \alpha_{f0} + \frac{1}{2} + 1 \right) \ln v_f \\ &\quad - \left( \beta_{f0} + \frac{1}{2} \left( \widehat{f}_{PM}^2 + \widehat{\Sigma}_{jj} \right) \right) v_f^{-1} \end{aligned} \tag{66}$$

from which we can establish the proportionality corresponding to  $q_4(v_f)$ :

$$q_{3j}(v_j) \propto v_j^{-(\alpha_{f0} + \frac{1}{2} + 1)} \exp \left\{ - \left[ \beta_{f0} + \frac{1}{2} \left( \widehat{f}_{PM}^2 + \widehat{\Sigma}_{jj} \right) \right] v_j^{-1} \right\} \tag{67}$$

Equation (67) leads to the following.

**Intermediate conclusion 3.** The probability distribution function  $q_4(v_f)$  is an inverse gamma distribution, with the parameters  $\alpha_{fj}$  and  $\beta_{fj}$ :

$$q_{3j}(v_j) = \text{IG} \left( v_j | \alpha_{fj}, \beta_{fj} \right), \begin{cases} \alpha_{fj} = \alpha_{f0} + \frac{1}{2} \\ \beta_{fj} = \beta_{f0} + \frac{1}{2} \left( \widehat{f}_{PM}^2 + \widehat{\Sigma}_{jj} \right) \end{cases} \tag{68}$$

Expressions (51), (60), and (68) resume the distributions families and the corresponding parameters for  $q_1(f)$ ,  $q_{2i}(v_{\epsilon_i})$ ,  $i \in \{1, 2, \dots, N\}$  and  $q_{3j}(v_j)$ ,  $j \in \{1, 2, \dots, M\}$ . However, the parameters corresponding to the multivariate normal distribution are expressed via  $\widetilde{V}_\epsilon^{-1}$  and  $\widetilde{V}_f^{-1}$  (and by extension, all elements forming the three matrices  $\widetilde{v}_{\epsilon_i}^{-1}$ ,  $i \in \{1, 2, \dots, N\}$  and  $\widetilde{v}_{f_j}^{-1}$ ,  $j \in \{1, 2, \dots, M\}$ ).

• **Computation of  $\widetilde{V}_\epsilon^{-1}$ ,  $\widetilde{V}_f^{-1}$ :**

For an inverse gamma distribution with parameters  $\alpha$  and  $\beta$ ,  $\text{IG}(x|\alpha, \beta)$ , the following relation holds:

$$\langle x^{-1} \rangle_{\text{IG}(x|\alpha, \beta)} = \frac{\alpha}{\beta}$$

The prove of the above relation is done by direct computation, using the analytical expression of the inverse gamma distribution:

$$\begin{aligned} \langle x^{-1} \rangle_{\text{IG}(x|\alpha, \beta)} &= \int x^{-1} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp \left\{ -\frac{\beta}{x} \right\} dx \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\beta^{\alpha+1}} \int \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} x^{-(\alpha+1)-1} \\ &\quad \exp \left\{ -\frac{\beta}{x} \right\} dx = \\ &= \frac{\alpha}{\beta} \underbrace{\int \text{IG}(x|\alpha+1, \beta) dx}_1 = \frac{\alpha}{\beta} \end{aligned}$$

Since  $q_{2i}(v_{\epsilon_i})$ ,  $i \in \{1, 2, \dots, N\}$  and  $q_{3j}(v_j)$ ,  $j \in \{1, 2, \dots, M\}$  are inverse gamma distributions, with parameters  $\alpha_{\epsilon_i}$  and  $\beta_{\epsilon_i}$ ,  $i \in \{1, 2, \dots, N\}$ , respectively,  $\alpha_{fj}$  and  $\beta_{fj}$ ,  $j \in \{1, 2, \dots, M\}$ , we can express the expectancies  $\widetilde{v}_{\epsilon_i}^{-1}$  and  $\widetilde{v}_{f_j}^{-1}$  via the parameters of the two inverse gamma distributions using the result above:

$$\widetilde{v}_{\epsilon_i}^{-1} = \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} ; \quad \widetilde{v}_{f_j}^{-1} = \frac{\alpha_{fj}}{\beta_{fj}} \tag{69}$$

Using the notation introduced in (37) and (39), we obtain:

$$\begin{aligned} \widetilde{V}_\epsilon^{-1} &= \begin{bmatrix} \frac{\alpha_{\epsilon_1}}{\beta_{\epsilon_1}} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & \frac{\alpha_{\epsilon_N}}{\beta_{\epsilon_N}} \end{bmatrix} = \widehat{V}_\epsilon^{-1} ; \\ \widetilde{V}_f^{-1} &= \begin{bmatrix} \frac{\alpha_{f_1}}{\beta_{f_1}} & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & \frac{\alpha_{f_j}}{\beta_{f_j}} & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & \frac{\alpha_{f_M}}{\beta_{f_M}} \end{bmatrix} = \widehat{V}_f^{-1} \end{aligned} \tag{70}$$

*Remark.* In Eq. (70), we have introduced other notations for  $\widetilde{V}_\epsilon^{-1}$  and  $\widetilde{V}_f^{-1}$ . All three values were expressed during the model via unknown expectancies, but at this point, we arrive at expressions that do not contain any more integrals to be computed. Therefore, the new notations represent the final expressions for the density functions  $q$  that depend only on numerical hyperparameters, set in the prior modeling.

**Appendix 3**

**Computations for PM estimation via VBA, full separability**

This section presents the computation for the PM estimation, via VBA, full separability (Subsection 4.3). The expression of the logarithm  $\ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  was established in the preview section (Eq. (33)).

• **Expression of  $q_{1j}(f_j)$ :**

Using Eq. (41):

$$\begin{aligned} & \langle \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_{1-j}(f_j) q_2(v_\epsilon) q_3(v_f)} \\ &= C - \frac{1}{2} \left\langle \left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 \right\rangle_{q_{1-j}(f_j)} \\ & \quad - \frac{1}{2} \left\langle \left\| \left( \widetilde{\mathbf{V}}_f^{-1} \right)^{1/2} \mathbf{f} \right\|^2 \right\rangle_{q_{1-j}(f_j)} \end{aligned} \quad (71)$$

For the first norm, considering all the  $f_j$  free terms as constants, we have:

$$\begin{aligned} \left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 &= C + \left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} \mathbf{H}^j \right\|^2 f_j^2 \\ & \quad - 2\mathbf{H}^{jT} \widetilde{\mathbf{V}}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}^{-j} \mathbf{f}^{-j}) f_j \end{aligned} \quad (72)$$

where  $\mathbf{H}^j$  represents the column  $j$  of the matrix  $\mathbf{H}$ ,  $\mathbf{H}^{-j}$  represents the matrix  $\mathbf{H}$  except the column  $j$ , and  $\mathbf{f}^{-j}$  represents the vector  $\mathbf{f}$  except the element  $f_j$ . Introducing the notation

$$\widetilde{f}_k = \int f_k q_{1k}(f_k) df_k ; \quad \widetilde{\mathbf{f}}^{-j} = [\widetilde{f}_1 \dots \widetilde{f}_{j-1} \widetilde{f}_{j+1} \dots \widetilde{z}_M]^T \quad (73)$$

the expectancy of the first norm becomes:

$$\begin{aligned} \left\langle \left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} (\mathbf{g} - \mathbf{H}\mathbf{f}) \right\|^2 \right\rangle_{q_{1-j}(f_j)} &= C + \left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} \mathbf{H}^j \right\|^2 f_j^2 \\ & \quad - 2\mathbf{H}^{jT} \widetilde{\mathbf{V}}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}^{-j} \widetilde{\mathbf{f}}^{-j}) f_j \end{aligned} \quad (74)$$

The expectancy for the second norm, considering all the free  $f_j$  terms as constants:

$$\left\langle \left\| \left( \widetilde{\mathbf{V}}_f^{-1} \right)^{1/2} \mathbf{f} \right\|^2 \right\rangle_{q_{1-j}(f_j)} = C + \widetilde{v}_{f_j}^{-1} f_j^2 \quad (75)$$

From Eqs. (31) and (71) and Eqs. (74) and (75), the proportionality for  $q_{1j}(f_j)$  becomes:

$$\begin{aligned} q_{1j}(f_j) \propto \exp \left\{ \left( \left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} \mathbf{H}^j \right\|^2 + \widetilde{v}_{f_j}^{-1} \right) f_j^2 \right. \\ \left. - 2\mathbf{H}^{jT} \widetilde{\mathbf{V}}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}^{-j} \widetilde{\mathbf{f}}^{-j}) f_j \right\} \end{aligned} \quad (76)$$

Defining the criterion  $J(f_j) = \left( \left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} \mathbf{H}^j \right\|^2 + \widetilde{v}_{f_j}^{-1} \right) f_j^2 - 2\mathbf{H}^{jT} \widetilde{\mathbf{V}}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}^{-j} \widetilde{\mathbf{f}}^{-j}) f_j$ , we arrive to the following.

**Intermediate conclusion 4.** The probability distribution function  $q_{1j}(f_j)$  is a normal distribution.

In order to compute the mean of the normal distribution, it is sufficient to compute the solution that minimizes the criterion  $J(f_j)$ :

$$\frac{\partial J(f_j)}{\partial f_j} = 0 \Leftrightarrow \widehat{f}_{j\text{PM}} = \frac{\mathbf{H}^{jT} \widetilde{\mathbf{V}}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}^{-j} \widetilde{\mathbf{f}}^{-j})}{\left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} \mathbf{H}^j \right\|^2 + \widetilde{v}_{f_j}^{-1}} \quad (77)$$

For the variance, we apply the same identification strategy as in the previous case, obtaining:

$$q_{1j}(f_j) = \mathcal{N} \left( f_j | \widehat{f}_{j\text{PM}}, \text{var}_j \right), \quad \begin{cases} \widehat{f}_{j\text{PM}} = \frac{\mathbf{H}^{jT} \widetilde{\mathbf{V}}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}^{-j} \widetilde{\mathbf{f}}^{-j})}{\left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} \mathbf{H}^j \right\|^2 + \widetilde{v}_{f_j}^{-1}} \\ \text{var}_j = \frac{1}{\left\| \left( \widetilde{\mathbf{V}}_\epsilon^{-1} \right)^{1/2} \mathbf{H}^j \right\|^2 + \widetilde{v}_{f_j}^{-1}} \end{cases} \quad (78)$$

• **Expression of  $q_{2i}(v_{\epsilon_i})$**

The proportionality relation corresponding to  $q_{2i}(v_{\epsilon_i})$  established in Eq. (31) refers to  $v_{\epsilon_i}$ , so in the expression of  $\ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$ , all the terms free of  $v_{\epsilon_i}$  can be regarded as constants:

$$\begin{aligned} \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) &= C - \left( \alpha_{\epsilon_0} + 1 + \frac{1}{2} \right) \ln v_{\epsilon_i} \\ & \quad - \left( \beta_{\epsilon_0} + \frac{1}{2} (\mathbf{g}_i - \mathbf{H}_i \mathbf{f}) \right) v_{\epsilon_i}^{-1} \end{aligned} \quad (79)$$

With the notation:

$$\langle \mathbf{f} \rangle_{q_1(\mathbf{f})} = [\widehat{f}_{1\text{PM}} \dots \widehat{f}_{j\text{PM}} \dots \widehat{f}_{M\text{PM}}]^T \stackrel{\text{Not}}{=} \widehat{\mathbf{f}}_{\text{PM}} \quad (80)$$

the expectancy of the logarithm becomes:

$$\begin{aligned} & \langle \ln p(\mathbf{f}, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_{1j}(f_j) q_{2-i}(v_{\epsilon_i}) q_3(v_f)} \\ &= C - \left( \alpha_{\epsilon_0} + 1 + \frac{1}{2} \right) \ln v_{\epsilon_i} \\ & \quad - \left( \beta_{\epsilon_0} + \frac{1}{2} \left[ \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T + (\mathbf{g}_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 \right] \right) v_{\epsilon_i}^{-1} \end{aligned} \quad (81)$$

and the proportionality relation for  $q_{2i}(v_{\epsilon_i})$  becomes:

$$\begin{aligned} q_{2i}(v_{\epsilon_i}) \propto v_{\epsilon_i}^{-(\alpha_{\epsilon_0} + 1 + \frac{1}{2})} \exp \left\{ - \left( \beta_{\epsilon_0} + \frac{1}{2} \left[ \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T \right. \right. \right. \\ \left. \left. \left. + (\mathbf{g}_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 \right] \right) v_{\epsilon_i}^{-1} \right\} \end{aligned} \quad (82)$$

Equation 82 leads to the following.

**Intermediate conclusion 5.** The probability distribution function  $q_{2i}(v_{\epsilon_i})$  is an inverse gamma distribution, with the parameters  $\alpha_{\epsilon_i}$  and  $\beta_{\epsilon_i}$ .

$$q_{2i}(v_{\epsilon_i}) = \mathcal{IG}(v_{\epsilon_i} | \alpha_{\epsilon_i}, \beta_{\epsilon_i}), \begin{cases} \alpha_{\epsilon_i} = \alpha_{\epsilon_0} + \frac{1}{2} \\ \beta_{\epsilon_i} = \beta_{\epsilon_0} + \frac{1}{2} \left[ \mathbf{H}_i \widehat{\Sigma} \mathbf{H}_i^T + (g_i - \mathbf{H}_i \widehat{\mathbf{f}}_{\text{PM}})^2 \right] \end{cases} \quad (83)$$

• **Expression of  $q_{3j}(v_{f_j})$**

The proportionality relation corresponding to  $q_{3j}(v_{f_j})$  established in Eq. (31) refers to  $v_{f_j}$ , so in the expression of  $\ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_{\mathbf{f}} | \mathbf{g})$ , all the terms free of  $v_{f_j}$  can be regarded as constants:

$$\ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_{\mathbf{f}} | \mathbf{g}) = C - \frac{1}{2} \ln v_{f_j} - \frac{1}{2} \left\langle f_j^2 \right\rangle_{q_{1j}(f_j)} v_{f_j}^{-1} - (\alpha_{f_0} + 1) \ln v_{f_j} - \beta_{f_0} v_{f_j}^{-1} \quad (84)$$

The integral of the logarithm:

$$(\ln p(\mathbf{f}, \mathbf{v}_{\epsilon}, \mathbf{v}_{\mathbf{f}} | \mathbf{g}))_{q_{1j}(f_j) q_{2j}(v_{\epsilon_j}) q_{3j}(v_{f_j})} = C - \left( \alpha_{f_0} + \frac{1}{2} + 1 \right) \ln v_{f_j} - \left[ \beta_{f_0} + \frac{1}{2} (\widehat{f}_{j\text{PM}}^2 + \text{var}_j) \right] v_{f_j}^{-1} \quad (85)$$

Equation 85 leads to the following.

**Intermediate conclusion 6.** The probability distribution function  $q_{3j}(v_{f_j})$  is an inverse gamma distribution, with the parameters  $\alpha_{f_j}$  and  $\beta_{f_j}$ .

$$q_{3j}(v_{f_j}) = \mathcal{IG}(v_{f_j} | \alpha_{f_j}, \beta_{f_j}), \begin{cases} \alpha_{f_j} = \alpha_{f_0} + \frac{1}{2} \\ \beta_{f_j} = \beta_{f_0} + \frac{1}{2} (\widehat{f}_{j\text{PM}}^2 + \text{var}_j) \end{cases} \quad (86)$$

**Appendix 4**

**List of symbols and abbreviations**

*List of symbols*

During the article, all the terms written in bold represent vectors or matrices.

1. **H**—the matrix used in the linear model considered during all the article.  $\mathbf{H} \in \mathcal{M}_{N \times M}$ . The matrix corresponds to the IFT and can be derived from Eq. (2).
2. **H<sub>i</sub>** represents the *i* line of the matrix **H**.  $\mathbf{H}_i \in \mathcal{M}_{1 \times M}$
3. **g<sub>0</sub>** represents the “theoretical” signal, i.e., the signal corresponding to the considered model (2) that does not account for the noise,  $\mathbf{g}_0 = \mathbf{H}\mathbf{f}$ . During the synthetic simulation section, the comparison between the estimated signal  $\widehat{\mathbf{g}}_0$  and the theoretical signal **g<sub>0</sub>** is particular important, measuring if the propose algorithm selects the solution corresponding to the biological phenomena.

4. **f** represents the PC vector,  $\mathbf{f} \in \mathcal{M}_{1 \times M}$ . This is the fundamental unknown of our model. All the estimates of the PC vector are denoted  $\widehat{\mathbf{f}}$  and in specific cases the particular estimation used in the model is indicated:  $\widehat{\mathbf{f}}_{\text{JMAP}}$  or  $\widehat{\mathbf{f}}_{\text{PM}}$ . During the article, the subscript used for indicating an element of the PC vector is *i*:  $f_i$  and the element is not bold, being a scalar.
5. **ε** represents the errors:  $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \dots, \epsilon_N]^T \in \mathcal{M}_{N \times 1}$ , is an *N*-dimensional vector

*List of abbreviations*

1. CT—circadian time
2. CTS—circadian timing system
3. FFT—fast Fourier transform
4. IGSM—infinite Gaussian scale mixture
5. IP—inverse problem
6. JMAP—joint maximum a posteriori
7. KL—Kullback-Leibler
8. PC vector—periodic component vector
9. PM—posterior mean
10. RT-BIO—RealTime Biolumicorder
11. TSVD—truncated single value decomposition
12. TRM—Tikhonov regularization methods
13. VBA—variational Bayesian approximation
14. ZT—Zeitgeber time

**Additional file**

**Additional file 1: Synthetic case simulations: 10 dB and 15 dB** (PDF 309 kb)

**Competing interests**

The authors declare that they have no competing interests.

**Acknowledgements**

The authors wish to gratefully acknowledge the reviewers for critically reading the manuscript and suggesting substantial improvements.

**Author details**

<sup>1</sup>Laboratoire des signaux et systèmes (L2S), UMR 8506 CNRS–CentraleSupélec–Univ. Paris-Sud, CentraleSupélec, Plateau de Moulon, 91192 Gif-sur-Yvette, France. <sup>2</sup>Rythmes Biologiques et Cancers (RBC), UMR 776 INSERM–Univ. Paris-Sud, Campus CNRS, 94801 Villejuif, France. <sup>3</sup>Department of Molecular Biotechnology and Health Sciences, University of Turin, 10126 Turin, Italy.

Received: 15 February 2015 Accepted: 7 December 2015

Published online: 20 January 2016

## References

- ME Hastings, AB Reddy, ES Maywood, A clockwork web: circadian timing in brain and periphery, in health and disease. *Nat Rev Neurosci.* **4**(8), 649–661 (2008)
- C Saini, DM Suter, A Liani, P Gos, U Schibler, The mammalian circadian timing system: synchronization of peripheral clocks. *Cold Spring Harb Symp Quant Biol.* **76**, 39–47 (2011)
- C Dibner, U Schibler, U Albrecht, The mammalian circadian timing system: organisation and coordination of central and peripheral clocks. *Annu Rev Physiol.* **72**, 517–549 (2010)
- JA Mohawk, CB Green, JS Takahashi, Central and peripheral circadian clocks in mammals. *Annu Rev Neurosci.* **35**, 445–462 (2012)
- RE Mistlberger, DJ Skene, Social influences on mammalian circadian rhythms: animal and human studies. *Biol Rev.* **79**, 533–556 (2004)
- Y Toutou, E Haus (eds.), *Biologic rhythms in clinical and laboratory medicine* (Springer Berlin Heidelberg, Berlin, 1993). doi: 10.1007/978-3-642-78734-8
- X-M Li, A Mohammad-Djafari, M Dumitru, S Dulong, E Filipki, S Siffroi-Fernandez, A Mteyrek, F Scaglione, C Guettier, F Delaunay, F Levi, A circadian clock transcription model for the personalization of cancer chronotherapy. *Cancer Res.* **73**(24), 7176–7188 (2013)
- PF Innominato, S Giacchetti, T Moreau, R Smaaland, C Focan, GA Bjarnason, Prediction of survival by neutropenia according to delivery schedule of oxaliplatin-5-fluorouracil-leucovorin for metastatic colorectal cancer in a randomized international trial. *Chronobiol Int.* **28**, 586–600 (2011)
- PF Innominato, S Giacchetti, GA Bjarnason, C Focan, C Garufi, B Coudert, Prediction of overall survival through circadian rest-activity monitoring during chemotherapy for metastatic colorectal cancer. *Int J Cancer.* **131**, 2684–2692 (2012)
- E Ortiz-Tudela, A Mteyrek, A Ballesta, PF Innominato, F Levi, Cancer chronotherapeutics: experimental, theoretical, and clinical aspects. *Handb Exp Pharmacol.* **217**, 261–288 (2013)
- F Levi, A Okyar, S Dulong, PF Innominato, J Clairambault, Circadian timing in cancer treatment. *Ann Rev Pharmacol Toxicol.* **50**, 377–421 (2010)
- F Levi, U Schibler, Circadian rhythms: mechanisms and therapeutic implications. *Annu Rev Pharmacol Toxicol.* **47**, 593–628 (2007)
- L F, Circadian chronotherapy for human cancers. *Lancet Oncol.* **2**, 307–315 (2001)
- MC Mormont, F Levi, Cancer chronotherapy: principles, applications, and perspectives. *Cancer.* **97**, 155–169 (2003)
- E Filipki, VM King, X Li, TG Granda, MC Mormont, Host circadian clock as a control point in tumor progression. *J Natl Cancer Inst.* **94**, 690–697 (2002)
- L Fu, CC Lee, The circadian clock: pacemaker and tumor suppressor. *Nat Rev Cancer.* **3**, 350–361 (2003)
- PF Innominato, C Focan, T Gorlia, T Moreau, C Garufi, Circadian rhythm in rest and activity: a biological correlate of quality of life and a predictor of survival in patients with metastatic colorectal cancer. *Cancer Res.* **69**, 4700–4707 (2009)
- MC Mormont, J Waterhouse, P Bleuzen, S Giacchetti, A Jami, Marked 24-h rest/activity rhythms are associated with better quality of life, better response, and longer survival in patients with metastatic colorectal cancer and good performance status. *Clin Cancer Res.* **6**, 3038–3045 (2000)
- SE Sephton, RM Sapolsky, HC Kraemer, D Spiegel, Diurnal cortisol rhythm as a predictor of breast cancer survival. *J Natl Cancer Inst.* **92**, 994–1000 (2000)
- E Filipki, PF Innominato, M Wu, XM Li, S Iacobelli, Effects of light and food schedules on liver and tumor molecular clocks in mice. *J Nat Cancer Inst.* **97**, 507–517 (2005)
- A Okyar, E Piccolo, C Ahowesso, E Filipki, V Hossard, C Guettier, R La Sorda, N Tinari, S Iacobelli, F Levi, Strain- and sex-dependent circadian changes in *abcc2* transporter expression: implications for irinotecan chronotolerance in mouse ileum. *PLoS One.* **6**(6), e20393 (2011)
- C Saini, A Liani, T Curie, P Gos, F Kreppel, Y Emmenegger, L Bonacina, JP Wolf, YA Poget, P Franken, U Schibler, Real-time recording of circadian liver gene expression in freely moving mice reveals the phase-setting behavior of hepatocyte clocks. *Genes Dev.* **27**, 1526–1536 (2013)
- Gu L, WM Tsark, DA Brown, S Blanchard, TW Synold, SE Kane, A new model for studying tissue-specific *mdr1a* gene expression in vivo by live imaging. *Proc Natl Acad Sci USA.* **106**, 5394–5399 (2009)
- K Aczél, I Vajk, Separation of periodic and aperiodic sound components by employing frequency estimation, (2008). <http://www.scopus.com/inward/record.url?eid=2-s2.0-84863771902&partnerID=40&md5=e489960c7cb3987fa63da10c5e76317e>
- JU Blackford, RM Salomon, NG Waller, Detecting change in biological rhythms: a multivariate permutation test approach to fourier-transformed data. *Chronobiol Int.* **26**(2), 258–281 (2009)
- S Bourguignon, H Carfantan, New methods for fitting multiple sinusoids from irregularly sampled data. *Stat Methodol.* **5**(4), 318–327 (2008)
- MJ Costa, B Finkenstadt, V Roche, F Levi, PD Gould, J Foreman, K Halliday, A Hall, DA Rand, Inference on periodicity of circadian time series. *Biostatistics.* **14**(4), 792–806 (2013)
- CR Smith, GJ Erickson, PO Neudorfer. Parameter estimation in chirped signals, (1989), pp. 538–539. <http://www.scopus.com/inward/record.url?eid=2-s2.0-0024927097&partnerID=40&md5=daa8c2dd6a477336001496f89e296426>
- R Llinares, J Igual, J Miro-Borras, A Camacho. Atrial activity estimation using periodic component analysis, (2010). <http://www.scopus.com/inward/record.url?eid=2-s2.0-79959472102&partnerID=40&md5=4d3039a80f7bb1ea568085f07f93b0fb>
- BA Rosa, Y Jiao, S Oh, BL Montgomery, W Qin, J Chen, Frequency-based time-series gene expression repositioning using prism. *BMC Syst Biol.* **6**(1), 1–16 (2012). ISSN: 1752-0509, doi: 10.1186/1752-0509-6-69, <http://dx.doi.org/10.1186/1752-0509-6-69>
- K Paraschakis, R Dahlhaus, Frequency and phase estimation in time series with quasi periodic components. *J Time Ser Anal.* **33**(1), 13–31 (2012)
- Y Jiang, Y He, High-accuracy phase difference estimation between same frequency components in two periodic signals. *Diangong Jishu Xuebao/Trans China Electromagn Soc.* **21**(11), 116–120 (2006)
- N Radde, L Kaderali, Inference of an oscillating model for the yeast cell cycle. *Discret Appl Math.* **157**(10), 2285–2295 (2009)
- W Gersch, in *Decision and Control, 1987. 26th IEEE Conference on. Some applications of smoothness priors in time series*, vol. 26, (1987), pp. 1684–1689. doi: 10.1109/CDC.1987.272756, <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&number=4049585&isnumber=4049208>
- A Aderhold, D Husmeier, M Grzegorzczak, Statistical inference of regulatory networks for circadian regulation. *Stat Appl Genet Mol Biol.* **13**(3), 227–273 (2014)
- LA Walls, A Bendell, Time series methods in reliability. *Reliab Eng.* **18**(4), 239–265 (1987)
- P Babu, P Stoica. Sparse spectral-line estimation for nonuniformly sampled multivariate time series: SPICE, LKES and MSBL, (2012), pp. 445–449. <http://www.scopus.com/inward/record.url?eid=2-s2.0-84869834781&partnerID=40&md5=3813f3549f9b7eaa2228597d512c4658>
- A Deckard, RC Anafi, JB Hogenesch, SB Haase, J Harer, Design and analysis of large-scale biological rhythm studies: a comparison of algorithms for detecting periodic signals in biological data. *Bioinformatics (Oxford, England).* **29**(24), 3174–3180 (2013)
- L Holmstrom, I Launonen, Posterior singular spectrum analysis. *Stat Anal Data Min.* **6**(5), 387–402 (2013)
- J Hong, On modeling nonstationary geomagnetic signal, vol. 2, (1998), pp. 1593–1596. <http://www.scopus.com/inward/record.url?eid=2-s2.0-0032277163&partnerID=40&md5=46bac254ef8f7ebc13aea42a3c55468>
- J Zhang, Y Li, Y Zhu, B Li. Estimation and prediction of noise power based on variational Bayesian and adaptive ARMA time series, vol. 9159, (2014). <http://www.scopus.com/inward/record.url?eid=2-s2.0-84902290799&partnerID=40&md5=8f9593133a7e69bea5277f98b622e2e3>
- M West. Time-frequency decompositions: Bayesian model-based approaches, vol. 1, (1998), p. 276. <http://www.scopus.com/inward/record.url?eid=2-s2.0-0032268549&partnerID=40&md5=e90265c0f5e0a3b470de1c86981184fc>
- CK Carter, R Kohn, Semiparametric bayesian inference for time series with mixed spectra. *JR Stat Soc Ser B Stat Methodol.* **59**(1), 255–268 (1997)
- R Paroli, L Spezia, Bayesian inference in non-homogeneous markov mixtures of periodic autoregressions with state-dependent exogenous variables. *Comput Stat Data Anal.* **52**(5), 2311–2330 (2008)
- EJ McCoy, DA Stephens, Bayesian time series analysis of periodic behaviour and spectral structure. *Int J Forecast.* **20**(4), 713–730 (2004)



46. ER Morrissey, MA Juarez, KJ Denby, NJ Burroughs, Inferring the time-invariant topology of a nonlinear sparse gene regulatory network using fully bayesian spline autoregression. *Biostatistics*. **12**(4), 682–694 (2011)
47. AS Dabrye, Bayesian estimation for a poisson process with a discontinuous intensity [estimation bayesienne pour un processus de poisson d'intensite discontinue]. *Comptes Rendus de l'Academie des Sciences - Series I: Mathematics*. **328**(5), 427–430 (1999)
48. E Granqvist, GE Oldroyd, RJ Morris, Automated bayesian model development for frequency detection in biological time series. *BMC Syst Biol*. **5**(1), 1–4 (2011). ISSN: 1752-0509, doi: 10.1186/1752-0509-5-97, <http://dx.doi.org/10.1186/1752-0509-5-97>
49. G Padmanabhan, AR Rao, Maximum entropy spectra of some rainfall and river flow time series from southern and central india. *Theor Appl Climatol*. **37**(1–2), 63–73 (1986)
50. PA Sturrock, JD Scargle, A bayesian assessment of p-values for significance estimation of power spectra and an alternative procedure, with application to solar neutrino data. *Astron J*. **706**(1), 393–398 (2009)
51. G Huerta, M West, Bayesian inference on periodicities and component spectral structure in time series. *J Time Ser Anal*. **20**(4), 401–416 (1999)
52. G Demoment, A Houacine, A Herment, I Mouttappa. Adaptive Bayesian spectrum estimation, (1988), pp. 33–38. <http://www.scopus.com/inward/record.url?eid=2-s2.0-0024143384&partnerID=40&md5=bc0098500514d7c0c651609ee516db28>
53. DV Divine, F Godtlielsen, Bayesian modeling and significant features exploration in wavelet power spectra. *Nonlinear Process Geophys*. **14**(1), 79–88 (2007)
54. A Mohammad-Djafari, Bayesian approach with prior models which enforce sparsity in signal and image processing. *EURASIP J Adv Signal Process*. **1**, 52–71 (2012)
55. MJ Wainwright, EP Simoncelli, in *Adv Neural Inform Process Syst (NIPS'99)*, ed. by SA Solla, TK Leen, and K-R Müller. Scale mixtures of gaussians and the statistics of natural images, vol. 12 (MIT Press, Cambridge, MA, 2000), pp. 855–861. <http://papers.nips.cc/paper/1750-scale-mixtures-of-gaussians-and-the-statistics-of-naturalimages>
56. M Dumitru, A Mohammad-Djafari, in *the 34th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt 2014) Proceedings*. Estimating the periodic components of a biomedical signal through inverse problem modelling and Bayesian inference with sparsity enforcing prior, vol. 1641 (Château Clos Lucé, Parc Leonardo DaVinci, Amboise, France, 2014), pp. 548–555. doi:<http://dx.doi.org/10.1063/1.4906021>

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---