



# A Bayesian approach for periodic components estimation for chronobiological signals

Mircea Dumitru

► **To cite this version:**

Mircea Dumitru. A Bayesian approach for periodic components estimation for chronobiological signals. Probability [math.PR]. Université Paris-Saclay, 2016. English. <NNT : 2016SACLS104>. <tel-01318048>

**HAL Id: tel-01318048**

**<https://tel.archives-ouvertes.fr/tel-01318048>**

Submitted on 19 May 2016

**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.

NNT : 2016SACLS104

THÈSE DE DOCTORAT  
DE  
L'UNIVERSITÉ PARIS-SACLAY  
PRÉPARÉE À  
L'UNIVERSITÉ PARIS-SUD

LABORATOIRE DES SIGNAUX ET SYSTÈMES

ECOLE DOCTORALE N° 580 STIC  
Sciences et technologies de l'information et de la communication

Traitement du Signal et des Images

Par

**M. Mircea Dumitru**

Approche bayésienne de l'estimation des composantes périodiques des signaux en  
chronobiologie

**Thèse présentée et soutenue à CentraleSupélec, le 25 Mars 2016 :**

**Composition du Jury :**

M. Kieffer, Michel  
M. Giovannelli, Jean-François  
M. Kuruoglu, Ercan Engin  
M. Renaux, Alexandre  
M. Djafari, Ali-Mohammad  
M. Lévi, Francis

Professeur des Universités, Paris-Sud  
Professeur des Universités, Bordeaux  
Chercheur senior CNR, ISTI, Italie  
Maître de conférences, Paris-Sud  
Directeur de recherche CNRS  
Directeur de recherche CNRS

Président  
Rapporteur  
Rapporteur  
Examineur  
Directeur de thèse  
Co-directeur de thèse



Thèse préparée au  
**Laboratoire des Signaux et Systèmes**  
*UMR 8506 (Université Paris-Sud - CNRS- Supélec)*  
Supélec, 3 rue Joliot-Curie  
91192 Gif-sur-Yvette, France



---

## APPROCHE BAYÉSIENNE DE L'ESTIMATION DES COMPOSANTES PÉRIODIQUES DES SIGNAUX EN CHRONOBIOLOGIE

### Résumé

La toxicité et l'efficacité de plus de 30 agents anticancéreux présentent de très fortes variations en fonction du temps de dosage. Par conséquent, les biologistes qui étudient le rythme circadien ont besoin d'une méthode très précise pour estimer le vecteur de composantes périodiques (CP) de signaux chronobiologiques. En outre, dans les développements récents, non seulement la période dominante ou le vecteur de CP présentent un intérêt crucial, mais aussi leurs stabilités ou variabilités. Dans les expériences effectuées en traitement du cancer, les signaux enregistrés correspondant à différentes phases de traitement sont courts, de sept jours pour le segment de synchronisation jusqu'à deux ou trois jours pour le segment après traitement. Lorsque'on étudie la stabilité de la période dominante nous devons considérer des signaux très court par rapport à la connaissance *a priori* de la période dominante, placée dans le domaine circadien. Les approches classiques fonées sur la transformée de Fourier (TF) sont inefficaces (i.e. manque de précision) compte tenu de la particularité des données (i.e. la courte longueur). Une autre particularité des signaux qui est prise en considération dans ces expériences, est le niveau de bruit. Ces signaux étant très bruités, il est difficile de déterminer les composantes périodiques associées aux phénomènes biologiques et de les distinguer de celles qui sont associées au bruit. Dans cette thèse, nous proposons une nouvelle méthode pour l'estimation du vecteur de CP des signaux biomédicaux, en utilisant les informations biologiques *a priori* et en considérant un modèle qui représente le bruit.

Les signaux enregistrés dans le cadre d'expériences développées pour le traitement du cancer ont un nombre limité de périodes. Cette information *a priori* peut être traduite comme la parcimonie du vecteur de CP. La méthode proposée considère l'estimation de vecteur de CP comme un problème inverse en utilisant l'inférence bayésienne générale afin de déduire toutes les inconnues de notre modèle, à savoir le vecteur de CP mais aussi les hyperparamètres (i.e. les variances associées). L'information *a priori* de parcimonie est modélisée en utilisant une loi *a priori* renforcement la parcimonie. Dans cette thèse, nous proposons une distribution de Student, considérée comme la distribution marginale d'une loi bivariée - la distribution Normale - Inverse Gamma. En fait, lorsque l'égalité entre les paramètres de forme et d'échelle, de la distribution Inverse Gamma n'est pas imposée, la marginale de la distribution Normale-Inverse Gamma est une généralisation de la distribution de Student. Nous construisons un modèle hiérarchique où nous attribuons aussi une loi *a priori* pour les hyperparamètres. L'expression de la loi conjointe *a posteriori* du vecteur de CP et des hyperparamètres est obtenue par la règle de Bayes et les inconnues sont estimées soit par Maximum A Posteriori (MAP) soit par l'espérance *a posteriori* (EAP). Pour le calcul de EAP, l'expression de la loi *a posteriori* est approchée par une loi séparables en utilisant l'approximation bayésienne variationnelle (ABV), via la divergence de Kullback-Leibler (KL). Deux possibilités sont envisagées : une approximation avec des lois partiellement séparables ou entièrement séparable. Ces algorithmes sont présentés en détail et sont comparés avec ceux correspondant au modèle gaussien. Nous examinons la convergence des algorithmes et donnons des résultats de simulation afin de comparer leurs performances. Enfin, nous montrons des résultats de simulation sur des données synthétiques et réelles dans une application de traitement du cancer. Les données réelles utilisées dans cette thèse représentent des modèles de repos-activité et d'expression des gènes de KI / KI Per2 : :luc souris luc, âgées de 10 semaines, seules dans leur cages des RT-BIO.

**Mots-clés** : Estimation de composantes périodiques, Problèmes inverses, Approches bayésiennes, Modèle hiérarchique, Renforcement de parcimonie, Student-t généralisée, chronobiologie, chronothérapie, Gènes de l'horloge, Rythme circadien, Traitement du cancer.

---

## Abstract

The toxicity and efficacy of more than 30 anticancer agents presents very high variations, depending on the dosing time. Therefore the biologists studying the circadian rhythm require a very precise method for estimating the Periodic Components (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 seven days for the synchronization segment to two or three 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 distinguish them from the ones associated with the noise is a difficult task. In this thesis 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 the 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 all the unknowns of our model, i.e. the PC vector but also the hyperparameters. The sparsity prior information is modelled using a sparsity enforcing prior law. In this thesis we propose a Student-t distribution, viewed as the marginal distribution of a bivariate Normal - Inverse Gamma distribution. In fact, when the equality between the shape and scale parameters corresponding to the Inverse Gamma distribution is not imposed, the marginal of the Normal-Inverse Gamma distribution is a generalization of the Student-t distribution. We build a general Infinite Gaussian Scale Mixture (IGSM) hierarchical model where we also assign prior distributions for the hyperparameters. The expression of the joint posterior law of the unknown PC vector and the hyperparameters is obtained via the 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. Two possibilities are considered: an approximation with partially separable distributions and an approximation with a fully separable one. The algorithms are presented in detail and are compared with the ones corresponding to the Gaussian model. We examine the practical 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 thesis examines the rest-activity patterns and gene expressions of KI/KI Per2::luc mouse, aged 10 weeks, singly housed in RT-BIO.

**Keywords** : Periodic Components (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..



## Acknowledgements

---

First and foremost I would like to express my deep gratitude for my advisor *Ali Mohammad-Djafari*. It has been an honour and a privilege to be one of his Ph.D. students. I am very grateful for his continuous support and patient guidance, for all the encouragements and advices he has provided throughout my time as his student and especially for his friendship.

The members of the Groupe Problèmes Inverses have contributed immensely to my personal and professional time at L2S. I am grateful to all the permanent members of the group *Aurelia Fraysse, Nicolas Gac, Mathieu Kowalski, François Orieux* and *Thomas Rodet* and to all the postdocs and PhD students: *Sha Zhu, Caifang Cai, Ning Chu, Yuling Zheng, Long Chen, Leila Gharshalli, Li Wang, Fengchen Fang, Alina Mereşescu* and *Thomas Boulay* for their friendship and for their help.

Also, I would like to express my gratitude for my advisor *Francis Lévi* and all the members of *Rythmes Biologiques et Cancers* laboratory *Xiao-Mei Li, Alexandre Arbaud, Jacques Beau, Sandrine Dulong, Virginie Hossard, Pasquale Innominato, Ali Mteyrek, Alper Okyar, Narin Öztürk, Véronique Roche, Elisabet Ortiz Tudela* and *Monique Lévi* for their warm reception in the their laboratory and for introducing me in the fascinating field of chronobiology.

For my dissertation I would like to thank my reading committee members *Ercan Engin Kuruoglu* and *Jean-François Giovannelli* and my examiners *Michel Kieffer* and *Alexandre Renaux*. It was an honour to have them as jury members and I am deeply grateful for all the time they invested for reading the manuscript and for their insightful comments and encouragement.

I would like to thank *John Skilling* for introducing me to nested sampling. Working with him was an honour.

My sincere thanks also goes to all my friends from CentraleSupélec: *Achal Agrawal, Laurence Antunes, Simona Baghai, Ejder Bastug, Anne Batalie, Sofiane Ben Chabane, Islam Boussaada, Yacine Bouzidi, Patrice Brault, Hugo Cavalera, Frédéric Desprez, Bakarime Diomande, Sette Diop, Bernard Duchêne, Mohamed Farouq, Marcel Geamănu, Maryvonne Giron, Franck Guimonet, Fethullah Güneş, Israel Hinostroza, Djawad Hamache, Trong-Bien Hoang, Dina Irofti, Céline Labrude, Mohammed-Tahar Laraba, Dominique Lesselier, Moussa Kefel, Rajesh Koduri, Mohammad Koteich, Wanchao Ma, Delphine Maucherat, Azza Mokadem, Panagiotis Piteros, Chengfang Ren, Karim Tadrict, Miassa-Amira Taleb, Mohammad Toufayli* and *Chao Zhang*.

Lastly, I would like to thank and dedicate this thesis to my sister, *Irina*.

This dissertation would not have been possible without funding from the ERASysBio+.





# Contents

<b>1</b>	<b>Résumé</b>	<b>13</b>
<b>2</b>	<b>Introduction</b>	<b>21</b>
2.1	Biological context: from chronobiology to chronotherapy . . . . .	22
2.2	Experiments description . . . . .	23
2.3	Chronobiological signals: particularities . . . . .	24
2.4	Summary description of the following chapters . . . . .	26
<b>3</b>	<b>Classical Methods</b>	<b>27</b>
3.1	Basic example and notations . . . . .	27
3.2	Fourier Transform . . . . .	29
3.3	Zero Padding . . . . .	31
3.4	Direct or Generalized Inversion . . . . .	32
3.5	Least Squares and regularized LS . . . . .	37
3.6	$L_1$ regularization and LASSO . . . . .	43
3.7	Conclusion and why go to Bayesian . . . . .	48
<b>4</b>	<b>Linear Model, Inverse Problem and Bayesian approach with sparsity enforcing prior</b>	<b>51</b>
4.1	Forward model and inverse problem approach . . . . .	52
4.2	General Bayesian inference . . . . .	55
4.3	Sparsity enforcing priors . . . . .	56
4.4	Generalized Student-t hierarchical model . . . . .	58
4.4.1	Generalized Student-t hierarchical prior model . . . . .	58
4.4.2	Joint MAP Estimation . . . . .	63
4.4.3	Posterior Mean (via VBA, partial separability) . . . . .	65
4.4.4	Prior Inverse Gamma parameters setting for $v_{f_j}$ . . . . .	72

4.4.5	Prior Inverse Gamma parameters setting for $v_{\epsilon_i}$	75
4.4.6	Posterior Mean (via VBA, full separability)	77
<b>5</b>	<b>Simulations: Synthetic data</b>	<b>79</b>
5.1	Simulations 05dB	80
5.1.1	Synthetic data 05dB	80
5.1.2	Data 05dB	81
5.1.3	JMAP IGSM 05dB	82
5.1.4	PM (via VBA, partial separability) IGSM 05dB	84
5.1.5	PM (via VBA, full separability) IGSM 05dB	85
5.1.6	Methods comparison 05dB	86
5.1.7	Error comparison 05dB	87
5.2	Synthetic data 10dB	91
5.2.1	Data 10dB	91
5.2.2	JMAP IGSM 10dB	92
5.2.3	PM (via VBA, partial separability) IGSM 10dB	93
5.2.4	PM (via VBA, full separability) IGSM 10dB	95
5.2.5	Methods comparison 10dB	96
5.2.6	Error comparison 10dB	97
5.3	Synthetic data 15dB	100
5.3.1	Data 15dB	100
5.3.2	JMAP IGSM 15dB	101
5.3.3	PM (via VBA, partial separability) IGSM 15dB	102
5.3.4	PM (via VBA, full separability) IGSM 15dB	104
5.3.5	Methods comparison 15dB	105
5.3.6	Error comparison 15dB	106
5.4	Student-t PM via VBA vs. LASSO 05dB	109
5.5	Student-t PM via VBA vs. LASSO 10dB	114
<b>6</b>	<b>Simulations: Real data</b>	<b>121</b>
6.1	CT 502 A1: Photon Absorption and Activity	122
6.1.1	Dominant Period evolution: Activity	123
6.1.2	Dominant Period evolution: Photon	127

6.1.3	Conclusions: Dominant Period evolution . . . . .	131
6.2	CT 502 A2: Photon Absorption and Activity . . . . .	132
6.2.1	Dominant Period evolution: Activity . . . . .	132
6.2.2	Dominant Period evolution: Photon . . . . .	137
6.2.3	Conclusions: Dominant Period evolution . . . . .	140
6.3	CT 502 A3: Photon Absorption and Activity . . . . .	142
6.3.1	Dominant Period evolution: Activity . . . . .	142
6.3.2	Dominant Period evolution: Photon . . . . .	146
6.3.3	Conclusions: Dominant Period evolution . . . . .	150
6.4	CT 502 B1: Photon Absorption and Activity . . . . .	151
6.4.1	Dominant Period evolution: Activity . . . . .	152
6.4.2	Dominant Period evolution: Photon . . . . .	156
6.4.3	Conclusions: Dominant Period evolution . . . . .	161
6.5	CT 502 B2: Photon Absorption and Activity . . . . .	161
6.5.1	Dominant Period evolution: Activity . . . . .	162
6.5.2	Dominant Period evolution: Photon . . . . .	164
6.6	CT 502 B3: Photon Absorption and Activity . . . . .	168
6.6.1	Dominant Period evolution: Activity . . . . .	169
6.6.2	Dominant Period evolution: Photon . . . . .	172
6.6.3	Conclusions: Dominant Period evolution . . . . .	177
<b>7</b>	<b>Conclusions and perspectives</b>	<b>179</b>
7.1	Conclusions . . . . .	179
7.2	Perspectives . . . . .	181
7.2.1	Multicomponent case: PC amplitudes vectors for multiple inputs . . . . .	181
7.2.2	Sparsity Rate: initialisation and week dependencies . . . . .	182
<b>A</b>	<b>Generalized Student-t</b>	<b>185</b>
A.1	Context . . . . .	185
A.2	Student-t: Origins . . . . .	187
A.2.1	ML parameter estimation . . . . .	187
A.2.2	Bayesian parameter estimation . . . . .	188
A.3	Generalization from the Bayesian point of view . . . . .	189

A.4 Sparsity Rate and Generalized Student-t . . . . .	190
A.5 Conclusion . . . . .	192
<b>B Publications</b>	<b>195</b>
Références . . . . .	197

# 1

## Résumé

LA toxicité et l'efficacité de plus de 30 agents anticancéreux présentent de très fortes variations en fonction du temps de dosage. Par conséquent, les biologistes qui étudient le rythme circadien ont besoin d'une méthode très précise pour estimer le vecteur de composantes périodiques (CP) de signaux chronobiologiques. En outre, dans les développements récents, non seulement la période dominante ou le vecteur de CP présentent un intérêt crucial, mais aussi leur stabilités ou variabilités.

Les données réelles utilisées dans cette thèse, montrées dans la figure (1.1), representent des modèles de repos-activité et d'expression des gènes de KI / KI Per2::luc souris luc, âgées de 10 semaines, seules dans leur cages des RT-BIO.

Dans les expériences effectuées en traitement du cancer, les signaux enregistrés correspondant à différentes phases de traitement sont courts, de sept jours pour le segment de synchronisation jusqu'à deux ou trois jours pour le segment après traitement, figure (1.2) (modèles de repos-activité) et figure (1.3) (modèles d'expression des gènes).

Lorsqu'on étudie la stabilité de la période dominante nous devons considérer des signaux très court par rapport à la connaissance *a priori* de la période dominante, placée dans le domaine circadien. Les approches classiques basées sur la transformée de Fourier (TF) sont inefficaces (i.e. manque de précision) compte tenu de la particularité des données (i.e. la courte longueur). Une autre particularité des signaux qui est prise en considération dans ces expériences, est le niveau de bruit. Ces signaux étant très bruités, il est difficile de déterminer les composantes périodiques associées aux phénomènes biologiques et de les distinguer de celle qui sont associées au bruit.

Dans cette thèse, nous proposons une nouvelle méthode pour l'estimation du vecteur de CP des signaux biomédicaux, en utilisant les informations biologiques *a priori* et en considérant un modèle qui représente le bruit,

$$g = H_1 f_1 + H_2 f_2 + \epsilon \quad (1.1)$$

Les signaux enregistrés dans le cadre d'expériences développées pour le traitement du

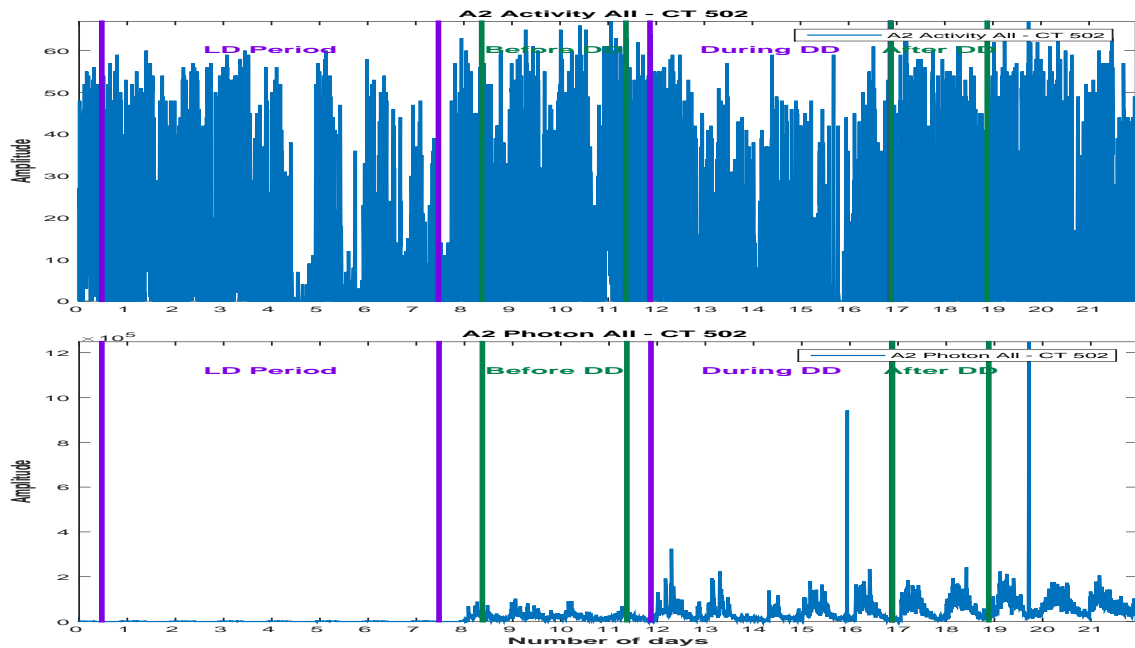


Figure 1.1: Exemple de signaux chronobiologiques obtenus dans des expériences de traitement du cancer: le signal repos-activité (en haut) et le signal l'expression des gènes (absorption de photons, en bas).

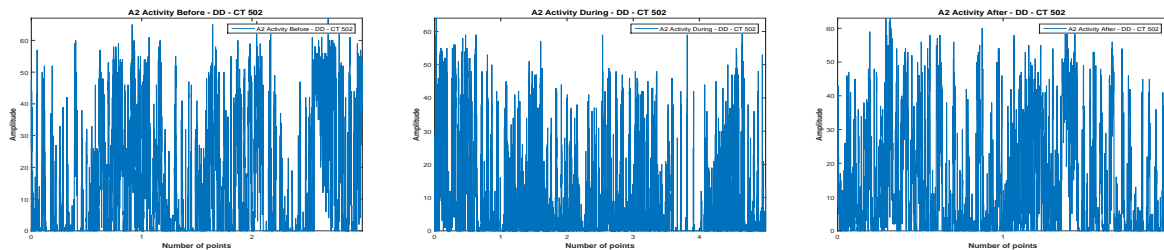


Figure 1.2: Les données brutes pour un signal represent des modelés de repos-activité en chronobiologie: avant le traitement (gauche, longueur 3 jours), pendant le traitement (centre, longueur de 5 jours), après le traitement (droite, longueur 2 jours).

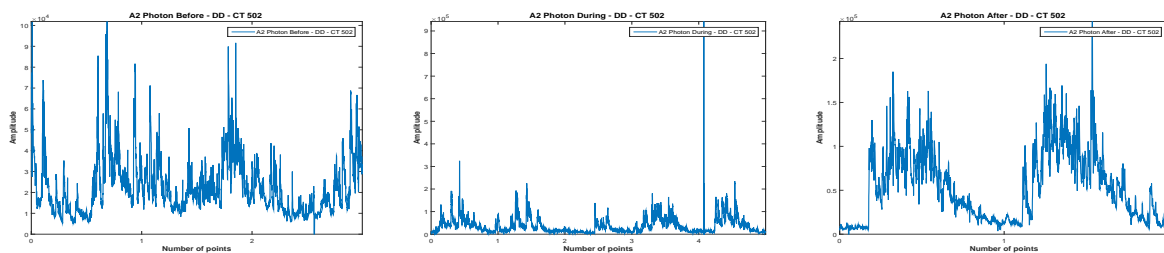


Figure 1.3: Les données brutes pour un signal represent l'expression des gènes en chronobiologie: avant le traitement (gauche, longueur 3 jours), pendant le traitement (centre, longueur de 5 jours), après le traitement (droite, longueur 2 jours).

cancer ont un nombre limité de périodes. Cette information *a priori* peut être traduit comme la parcimonie du vecteur de CP. La méthode proposée considère l'estimation de vecteur de CP comme un problème inverse en utilisant l'inférence bayésienne générale afin de déduire toutes les inconnues de notre modèle, à savoir le vecteur de CP mais aussi les hyperparamètres (i.e. les variances associées),

$$p(\mathbf{f}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f} | \boldsymbol{\theta}_2) p(\boldsymbol{\theta}_1) p(\boldsymbol{\theta}_2) \quad (1.2)$$

où  $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)$  représente les hyperparamètres qui apparaissent dans le modèle.

L'information *a priori* de parcimonie est modélisée en utilisant une loi *a priori* renforce la parcimonie. Dans cette thèse, nous proposons une distribution de Student, pour sa propriété de longue queue, présentée dans la figure (1.4).

$$St(f_{kj} | \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{f_{kj}^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad (1.3)$$

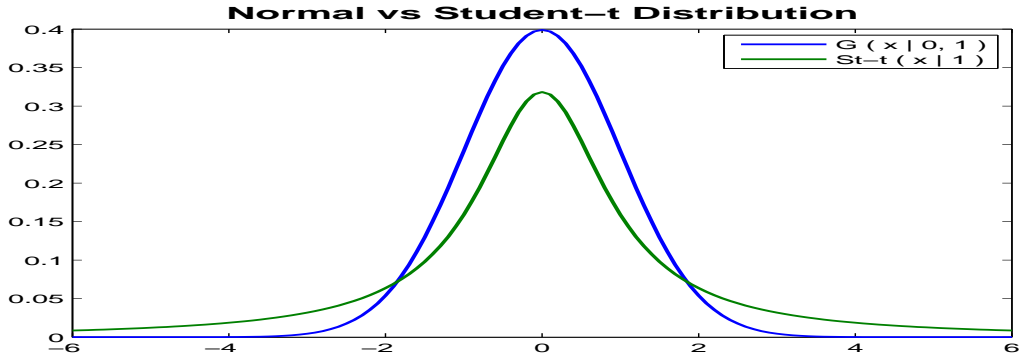


Figure 1.4: Densité de la loi Normale et densité de la loi de Student.

La distribution de Student est considérée comme la distribution marginale d'une loi bivariée - la distribution Normale - Inverse Gamma,

$$f_{kj} | \nu \sim \int \mathcal{N}(f_{kj} | 0, v_{f_j}) \mathcal{IG}(v_{f_j} | \frac{\nu}{2}, \frac{\nu}{2}), \quad k \in \{1, 2\}, \quad j \in \{1, 2, \dots, M\}. \quad (1.4)$$

En fait, lorsque l'égalité entre les paramètres de forme et d'échelle, de la distribution Inverse Gamma n'est pas imposée, on a

$$f_{kj} | \alpha_{f_0}, \beta_{f_0} \sim \int \mathcal{N}(f_{kj} | 0, v_{f_j}) \mathcal{IG}(v_{f_j} | \alpha_{f_0}, \beta_{f_0}), \quad k \in \{1, 2\}, \quad j \in \{1, 2, \dots, M\}. \quad (1.5)$$

La marginale de la distribution Normale-Inverse Gamma est une généralisation de la distribution de Student

$$p(f_{kj} | \alpha_{f_0}, \beta_{f_0}) = \frac{\Gamma(\alpha_{f_0} + \frac{1}{2})}{\sqrt{2\beta_{f_0}\pi} \Gamma(\alpha_{f_0})} \left(1 + \frac{f_{kj}^2}{2\beta_{f_0}}\right)^{-(\alpha_{f_0} + \frac{1}{2})} = St_g(f_{kj} | \alpha_{f_0}, \beta_{f_0}). \quad (1.6)$$



La distribution bivariée Normale - Inverse Gamma, modélisant  $(f_{kj}, v_{f_j})$  et la distribution marginale modélisant  $f_{kj}$  sont présentés dans la figure (1.5).

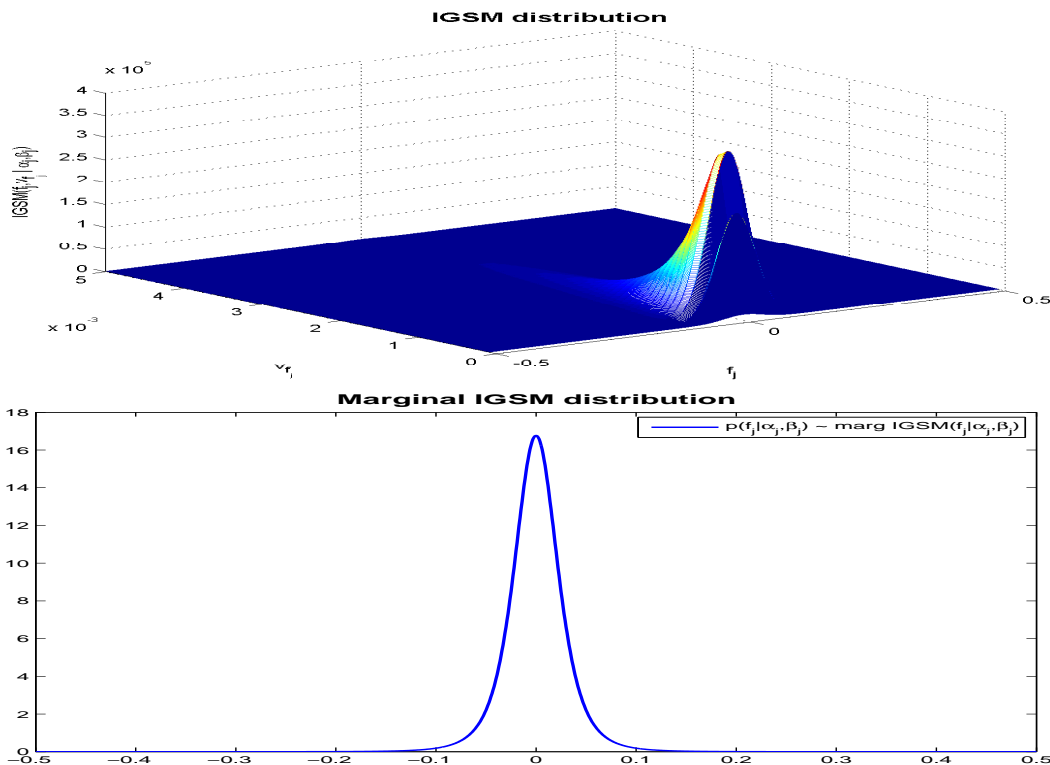


Figure 1.5: La distribution bivariée Normale - Inverse Gamma modélisant  $(f_{kj}, v_{f_j})$  et la distribution marginale modélisant  $f_{kj} | \alpha_{f_0}, \beta_{f_0}$ .

Nous construisons un modèle hiérarchique où nous attribuons aussi une loi *a priori* pour les hyperparamètres, Equation (1.7).

$$\begin{cases} p(\mathbf{g} | \mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon) = \mathcal{N}(\mathbf{g} | \mathbf{H}_1 \mathbf{f}_1 + \mathbf{H}_2 \mathbf{f}_2, \mathbf{V}_\epsilon) \\ p(\mathbf{v}_\epsilon | \alpha_{\epsilon 0}, \beta_{\epsilon 0}) = \prod_{i=1}^N \mathcal{IG}(v_{\epsilon_i} | \alpha_{\epsilon_i 0}, \beta_{\epsilon_i 0}) \\ p(\mathbf{f}_1 | \mathbf{v}_f) = \mathcal{N}(\mathbf{f}_1 | \mathbf{0}, \mathbf{V}_f) \\ p(\mathbf{f}_2 | \mathbf{v}_f) = \mathcal{N}(\mathbf{f}_2 | \mathbf{0}, \mathbf{V}_f) \\ p(\mathbf{v}_f | \alpha_{f_0}, \beta_{f_0}) = \prod_{j=1}^M \mathcal{IG}(v_{f_j} | \alpha_{f_0}, \beta_{f_0}) \end{cases} \quad (1.7)$$

La forme analytique du modèle hiérarchique écrite par proportionnalités est le suivant:

$$\begin{cases} p(\mathbf{g} | \mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon) \propto |\mathbf{V}_\epsilon|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \|\mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 \right\} \\ p(\mathbf{v}_\epsilon | \alpha_{\epsilon 0}, \beta_{\epsilon 0}) \propto \prod_{i=1}^N \frac{\beta_{\epsilon 0}^{\alpha_{\epsilon 0}}}{\Gamma(\alpha_{\epsilon 0})} v_{\epsilon_i}^{-(\alpha_{\epsilon 0}+1)} \exp \left\{ -\beta_{\epsilon 0} v_{\epsilon_i}^{-1} \right\} \\ p(\mathbf{f}_1 | \mathbf{v}_f) \propto |\mathbf{V}_f|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1\|^2 \right\} \\ p(\mathbf{f}_2 | \mathbf{v}_f) \propto |\mathbf{V}_f|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_2\|^2 \right\} \\ p(\mathbf{v}_f | \alpha_{f_0}, \beta_{f_0}) \propto \prod_{j=1}^M \frac{\beta_{f_0}^{\alpha_{f_0}}}{\Gamma(\alpha_{f_0})} v_{f_j}^{-(\alpha_{f_0}+1)} \exp \left\{ -\beta_{f_0} v_{f_j}^{-1} \right\} \end{cases} \quad (1.8)$$

L'expression de la loi conjointe *a posteriori* du vecteur de CP et des hyperparamètres est obtenue par la règle de Bayes:

$$p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon) p(\mathbf{v}_\epsilon | \alpha_{\epsilon 0}, \beta_{\epsilon 0}) p(\mathbf{f}_1 | \mathbf{v}_f) p(\mathbf{f}_2 | \mathbf{v}_f) p(\mathbf{v}_f | \alpha_{f 0}, \beta_{f 0}). \quad (1.9)$$

Les inconnues sont estimées soit par Joint Maximum A Posteriori (JMAP), soit par l'espérance *a posteriori* (EAP). L'estimateur du Joint Maximum A Posteriori s'obtient en calculant le mode de la loi *a posteriori*,

$$\left( \widehat{\mathbf{f}}_1, \widehat{\mathbf{f}}_2, \widehat{\mathbf{v}}_\epsilon, \widehat{\mathbf{v}}_f \right) = \underset{(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\arg \max} p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) = \underset{(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\arg \min} \mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f), \quad (1.10)$$

où le critère  $\mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)$  est défini comme suit:

$$\mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f) = -\ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}). \quad (1.11)$$

L'algorithme JMAP est présenté en détail dans la figure (1.6). Pour le calcul de EAP, l'expression de la loi *a posteriori*  $p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  est approchée par une loi séparables  $q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$ :

$$p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \approx q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) = q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f) \quad (1.12)$$

en utilisant l'approximation bayésienne variationnelle (ABV), via la divergence de Kullback-Leibler (KL), défini comme suit:

$$\begin{aligned} \text{KL}(q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) : p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})) &= \\ &= \iint \dots \int q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \ln \frac{q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})}{p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})} d\mathbf{f}_1 d\mathbf{f}_2 d\mathbf{v}_\epsilon d\mathbf{v}_f, \end{aligned} \quad (1.13)$$

où les notations suivantes sont utilisées:

$$q_3(\mathbf{v}_\epsilon) = \prod_{i=1}^N q_{3i}(v_{\epsilon_i}) ; \quad q_4(\mathbf{v}_f) = \prod_{j=1}^M q_{4j}(v_{f_j}) ; \quad d\mathbf{v}_\epsilon = \prod_{i=1}^N dv_{\epsilon_i} ; \quad d\mathbf{v}_f = \prod_{j=1}^M dv_{f_j} \quad (1.14)$$

La minimisation peut se faire via l'optimisation alternée, résultant les proportionnalités suivants:

$$\left\{ \begin{array}{l} q_1(\mathbf{f}_1) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f)} \right\} \\ q_2(\mathbf{f}_2) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f)} \right\} \\ q_{3i}(v_{\epsilon_i}) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_{3-i}(v_{\epsilon_i}) q_4(\mathbf{v}_f)} \right\}, \quad i \in \{1, 2, \dots, N\} \\ q_{4j}(v_{f_j}) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_{4-j}(v_{f_j})} \right\}, \quad j \in \{1, 2, \dots, M\}, \end{array} \right. \quad (1.15)$$

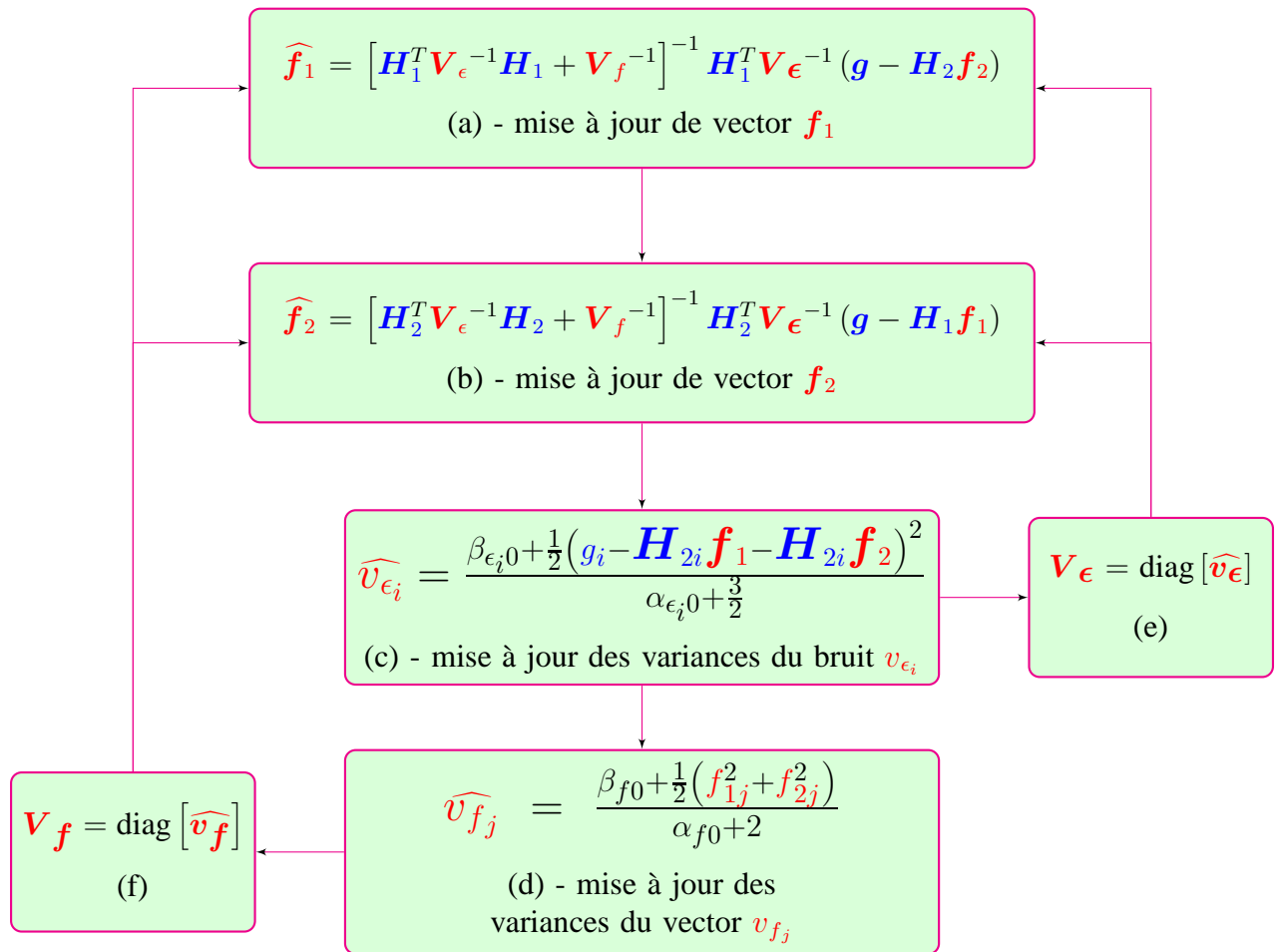


Figure 1.6: Les equations de mise à jour des paramètres pour l'estimation MAP.

en utilisant les notations:

$$q_{3-i}(v_{\epsilon_i}) = \prod_{k=1, k \neq i}^N q_{3k}(v_{\epsilon_k}) \quad ; \quad q_{4-j}(v_{f_j}) = \prod_{k=1, k \neq j}^M q_{4k}(v_{f_k}) \quad (1.16)$$

et aussi

$$\langle u(x) \rangle_{v(y)} = \int u(x) v(y) \, dy. \quad (1.17)$$

Deux possibilités sont envisagées: une approximation avec des lois partiellement séparables ou entièrement séparable. Ces deux algorithmes sont présentés en détail dans la figure (1.7) (approximation avec des lois partiellement séparables) et dans la figure (1.8) (approximation avec des lois entièrement séparable).

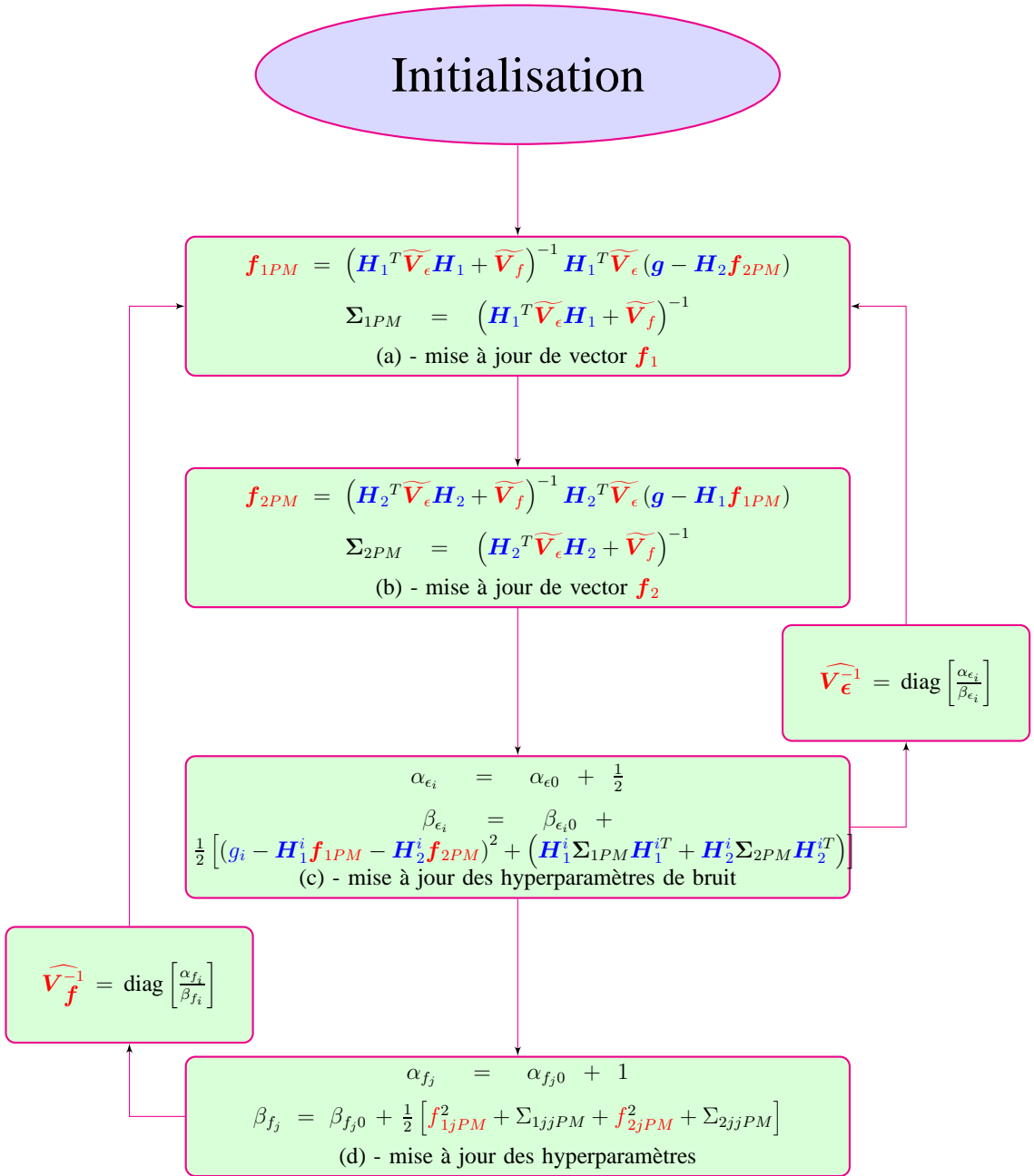


Figure 1.7: Les equations de mise à jour des paramètres pour l'estimation EAP via ABV, séparabilité partielle.

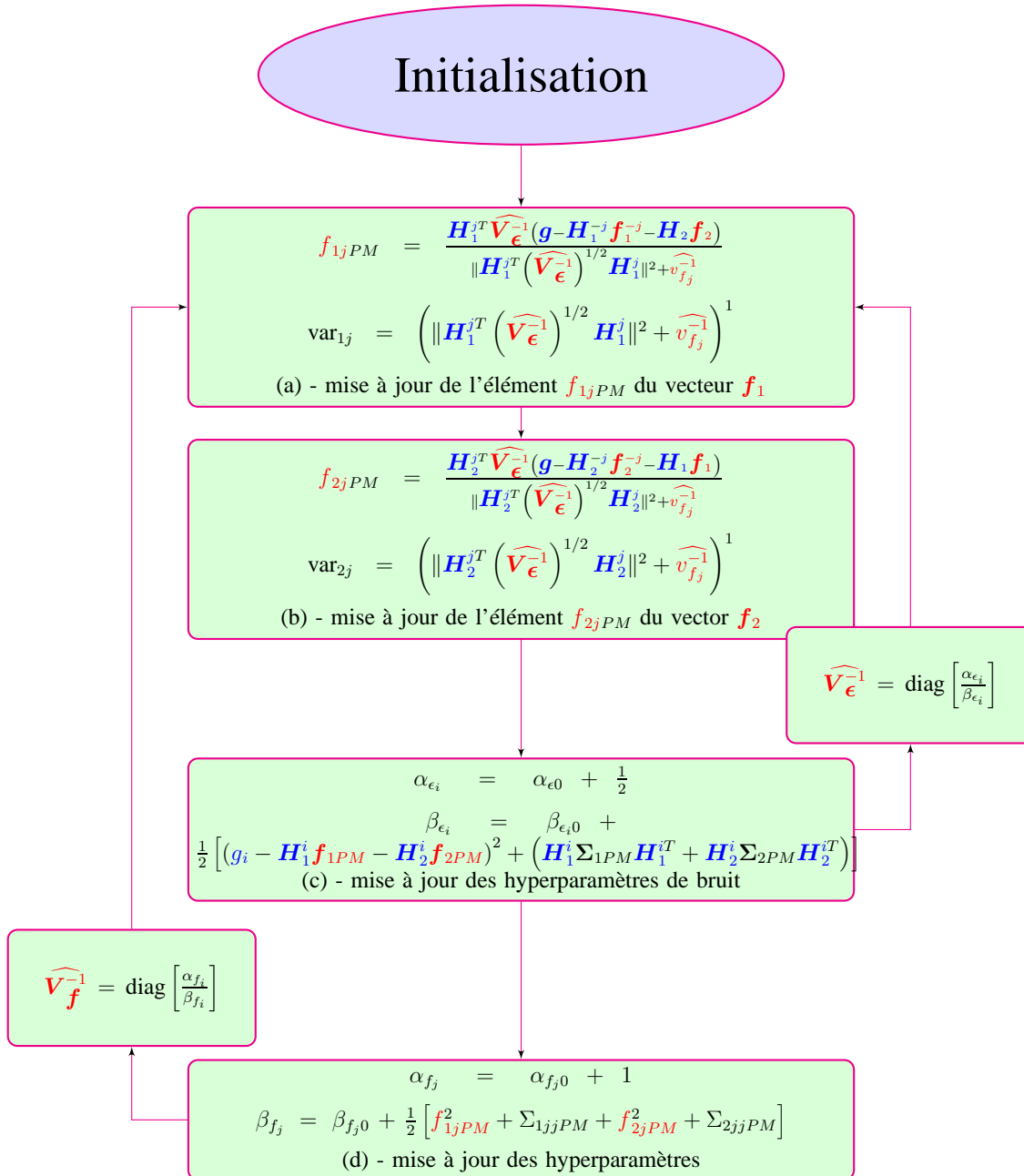


Figure 1.8: Les equations de mise à jour des paramètres pour l'estimation EAP via ABV, séparabilité totale.

# 2

## Introduction

### Contents

---

<b>2.1 Biological context: from chronobiology to chronotherapy</b> . . . . .	<b>22</b>
<b>2.2 Experiments description</b> . . . . .	<b>23</b>
<b>2.3 Chronobiological signals: particularities</b> . . . . .	<b>24</b>
<b>2.4 Summary description of the following chapters</b> . . . . .	<b>26</b>

---

**T**HIS first chapter gives a general presentation of the context of our work. From chronobiology and biological rhythms (in particular the circadian rhythms) to cancer treatment (chronotherapy), gene expressions (along with other circadian indicators, like temperature or rest-activity), signal processing and the particularities of such signals. The goal of this chapter is to give a short presentation of how the periodic components estimation of (short) time series arises naturally in the context of chronotherapy and cancer treatment. This chapter is organized as it follows:

In Section (2.1) we give a brief presentation of the biological context: we introduce the circadian rhythm, discuss its mechanism and its importance relative to cancer treatment.

Section (2.2) presents the experimental protocol for producing chronobiological signals: we give details of the kind of mice used in those experiments, of the clock-genes that are studied, discuss the protocol of the experiments and give brief details concerning the technology that allows recording such signals.

Finally, Section (2.3) makes the transition from chronobiology and cancer treatment to signal processing: we present such signals, we discuss their particularities and define our goal relative to such signals.

## 2.1 Biological context: from chronobiology to chronotherapy

---

Several biological processes in living organisms follow oscillations that repeat themselves about every 24 hours - these oscillations are called *circadian rhythms*. Together with other periodic phenomena, they are the object of study of chronobiology, [HRM08], [SSL<sup>+</sup>11] and [DSA10]. 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, [MGT12], [MS04] and [TH94]. 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/producing diffusible and neurophysiologic signals, [DSA10].

In mammals, the core of the molecular clock consists of about a dozen specific clock genes. The circadian oscillations in the single cells result from transcriptional and post-transcriptional activation and inhibition regulatory loops. Specifically, a positive branch, including the proteins Clock, Npas2 and Bmal1, and a negative branch, including Per (Per1 and Per2), CRY (Cry1 and Cry2) are interconnected by the nuclear orphan receptor Rev-erb $\alpha$ . The heterodimer of either Bmal1+Clock or Bmal1+Npas2 activate Per and Cry genes transcription. These heterodimers action is facilitated by Ezh2 (Polycomb protein), CK2 (casein kinase 2) and SIRT1. Per and Cry proteins accumulation and activity is also regulated by phosphorylation by protein kinases (CK $\delta$ ,  $\epsilon$ ), by ubiquitination (of Crys), by histone methylation and demethylation. Dec1 and Dec2 compete with the Bmal1+Clock/Npas2 heterodimers to bind Per and Cry genes, therefore reducing their activation. An accessory feedback loop sees the nuclear orphan receptors Rora $\alpha$ ,  $\beta$  and  $\gamma$  as activators of Bmal1 circadian transcription, whereas Rev-erb $\alpha$  and  $\beta$  inhibit it.

The period of the circadian timing system, which is *about* 24 hours, is therefore regularly calibrated by the succession of day and night (light and dark). It can be influenced by other environmental factors, such as socio-professional interactions and feeding times, [MS04]. The resulting circadian physiologic fluctuations are observed in sleep-wakefulness and rest-activity alternation, body temperature, cortisol secretion by the adrenal gland, melatonin secretion by the pineal gland, and they involve as well the sympathetic and the parasympathetic systems, [TH94].

Former studies have already shown how focusing on/taking chronobiology into account can improve anticancer treatments efficacy and reduce at the same time their toxicity/increase at the same time their tolerability, contrary to the previous "the worst the toxicity, the better the efficacy" paradigm, [LMDD<sup>+</sup>13], [IGM<sup>+</sup>11], [IGB<sup>+</sup>12] and [OTMB<sup>+</sup>13]. 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, [LOD<sup>+</sup>10].

Tolerability to anticancer treatments has been proven to depend significantly on their timing in respect to the circadian rhythms, measuring up to 10-fold changes in the tolerability to drug administration at different circadian times for 40 anticancer drugs in rodents, and up to 5-fold in patients, [LOD<sup>+</sup>10], [LS07]. Notably, chemotherapeutic agents proved to be at their best efficacy, both administered alone and combined, when they are also at their best tolerability level, that is to say when they are least toxic to the healthy tissues.

Furthermore, substantial interpatient variability of circadian rhythms have been observed and can be due to factors such as gender, age and genetic polymorphisms (amongst others), therefore anticancer drugs dosing and timing need to be personalized, at least for subgroups of patients with similar chronotoxicity (chronobiological) key features. Modulating drugs administration according to (the patient's) circadian rhythms is known as/called "chronotherapy", [F.01], [ML03]. 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, [FKL<sup>+</sup>02],[FL03],[IFG<sup>+</sup>09],[MWB<sup>+</sup>00],[SSKS00] and [FIW<sup>+</sup>05]. In order to optimize cancer treatment, once proven that a certain drug effects are susceptible to circadian rhythms, the best administration time must be identified. Furthermore, in case of interpatient variability in chronotoxicity, we want to personalize the standard chronotherapy to best fit to the patient's circadian specifics.

## 2.2 Experiments description

---

The experiments considered are focused on *irinotecan* and *everolimus*, two of the most common drugs used at the moment in cancer treatment. First, for each drug is proved the correlation with on the circadian rhythms in a rodent model, which has been proved to well represent the human circadian physiology, [LOD<sup>+</sup>10]. 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) using mice. Their circadian clock is synchronized by exposure to light for 12 hours, followed by 12 hours of dark, repeating this cycle for one week, and its rhythm is detected by measuring the expression of one or more of its core genes. The genes are *Bmal1*, *Per2*, *Rev-erb $\alpha$* .

Mice with a disrupted clock (clock-defective mice, obtained via the functional knock out of one of its genes, normally *Bmal1*, *Per2* or *Rev-erb $\alpha$* ) 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 the observed gene(s) are studied to observe whether the administration of the drug modifies them. Once defined the CTs or ZTs at which the drug best and worst tolerability is observed, we can look for the (specific tissutal) 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 [OPA<sup>+</sup>11].

Similarly, the experiments that we consider are studying everolimus toxicity in respect to the circadian expression of intestinal *Mdr1a*. Everolimus is an anticancer drug also used as immunosuppressant after transplants to prevent organ rejection and its pharmacokinetics also depend on *Mdr1a*, a P-glycoprotein (P-gp) which functions as a transmembrane drug transporter and is responsible for multidrug resistance in cancer cells that over-express it, [GP93] and [EIM90].

The circadian clocks of the mice used in such experiments are first synchronized to the same day-night alternation, where 12 hours of light are followed by 12 hours of dark (LD12:12). After synchronization, the mice are kept in constant darkness (DD), which implies to subtract the light as Zeitgeber, allowing the circadian clock to dictate their endogenous rhythm? and



normally the period of the circadian rhythm becomes shorter. After the start of DD, the mice are split in two groups: to one the drug is administered at the CT of the best chronotolerance, to the other at the CT of the worst chronotolerance. For each group one mouse is kept as control, no drug is administered to it. Throughout the experiment, gene expression and rest-activity pattern 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), [SLC<sup>+</sup>13]. 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 and is marked by bioluminescence measured by a photomultiplier tube. Common mouse strains used are C57BL/6-based [LMDD<sup>+</sup>13], [OPA<sup>+</sup>11] and 129S1/SvImJ, [LMB<sup>+</sup>09].

An exact estimation of the Periodic Components (PC) vector of the those chronobiological signals is needed.

## 2.3 Chronobiological signals: particularities

The major interest is the study of the periodicity of such chronobiological time series, i.e. the precise estimation of the PC vector and the stability of the dominant period, requiring a dominant periodic components variation analysis. The major limitation is given by their reduced length. Figure (2.1) presents the signals obtained in an experiment developed in

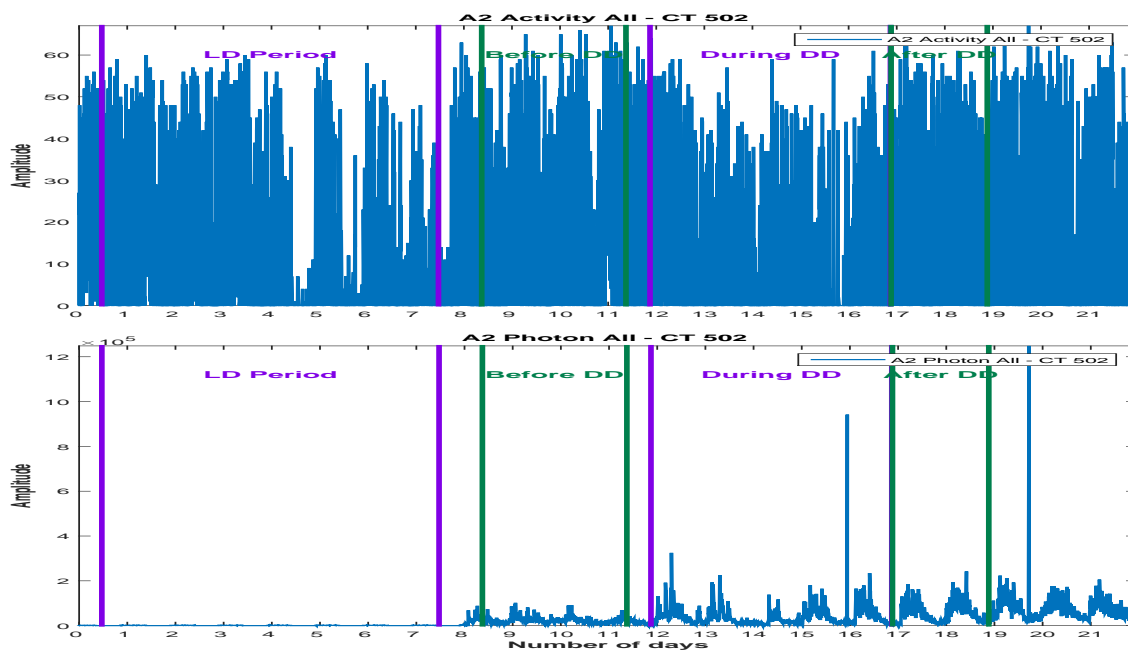


Figure 2.1: Example of chronobiological signals obtained in cancer treatment experiments: the rest-activity signal (up) and the gene expression signal (photon absorption, bottom)

the conditions described in Section (2.2). As mentioned, via the RT-BIO monitoring device, [SLC<sup>+</sup>13] the gene expression and the rest-activity patterns can be recorded. The experiment exemplified in Figure (2.1) corresponds to 22 days and both signals are sampled every minute (rest-activity pattern - up and gene expression - bottom). The gene expression is measured

by a photomultiplier tube, and the amplitude of the signal represents the photon absorption. For the rest-activity pattern, the locomotor activity is detected by an infrared sensor. For the PC vector estimation, four segments are of great interest: the first part, corresponding to the synchronization part, where 12 hours of light are followed by 12 hours of dark (LD12:12) and then the second part, in complete darkness (DD), which is divided in three segments, corresponding to before, during and after treatment. We note that the length of those segments vary between two or three days (the segments corresponding to the after treatment and to the before treatment) and 7 days (the segment corresponding to the synchronization, (LD12:12)). When studying the stability or the variability of the PC vector, even shorter signal must be considered. More precisely, for analysing the stability for the segment corresponding to the during treatment, at least 4 days length signals should be considered. In this context, we note a first particularity of the signal that are considered for the PC vector estimation. Relative to the prior knowledge, *an about 24 hours periodicity*, we deal with very short signals: three or four days length signals, relative to the an about one day period. Another particularity of such signals is represented by the noise. Both signals, the one corresponding to the gene expression and the one corresponding to the rest-activity pattern are (very) noisy signals due to the measurement errors.

The problem that is considered can be formulated as it follows: *precise estimation of the PC vector of noisy and very short signals relative to our knowledge for the principal period*. Examples of raw data signals for which the corresponding PC vector must be inferred are presented in Figure (2.2) and Figure (2.3) One fundamental particularity of such signals is

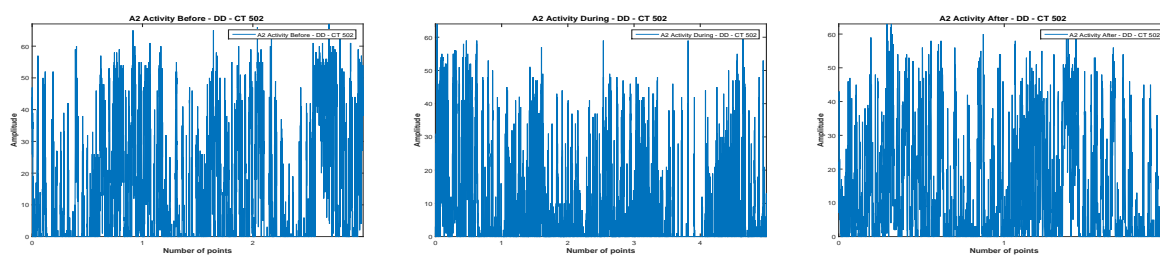


Figure 2.2: Raw data for a rest-activity pattern signal in chronobiology: before treatment (left, 3 days length), during treatment (center, 5 days length), after treatment (right, 2 days length)

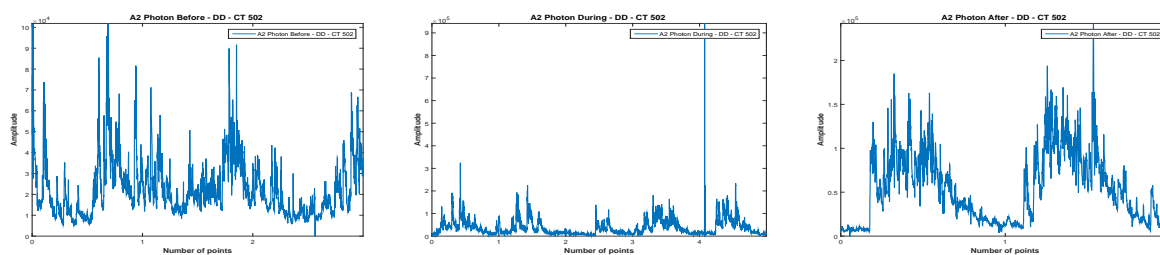


Figure 2.3: Raw data for a gene expression signal in chronobiology: before treatment (left, 3 days length), during treatment (center, 5 days length), after treatment (center, 2 days length)

given by the biological prior knowledge: the number of periodic components is small, i.e. the PC vector is sparse. In Chapter (3) we present the drawback of the classical methods, and in Chapter (4) we present the proposed method for the PC vector estimation, using the prior information.

### 2.4 Summary description of the following chapters

---

In Chapter (2) we have established the context of our work. Briefly, the fundamental goal of our work can be summarized as it follows:

***Estimate the sparse PC amplitudes vector corresponding to short duration signals relative to the (prior) dominant period with a given precision. The number of components is small (sparse PC amplitudes vector) but unknown.***

First, we present some classical methods that can be considered in order to address the above question. In Chapter (3) we examine the Fourier Transform techniques, the Direct or Generalized Inversion (DI), Least Squares (LS) and Regularized Least Squares techniques and finally we consider the  $L_1$  regularization and the LASSO method, the nearest state of art method to account for sparsity through  $L_1$  regularization term. We will show that given the particularities of the context, i.e. the short duration of the signals relative to the (prior) dominant period, the high level of noise in the signals, the above methods do not give satisfactory results.

In Chapter (4) we introduce the proposed solution, based on a forward model that accounts the noise, built in a similar way as the one corresponding to the Fourier transform but using the elements corresponding to the considered circadian PC amplitudes vector (in terms of limits and precision). The estimation of the PC amplitudes vector (but also the other hyperparameters involved in the model, i.e. the variances corresponding to the PC amplitudes vector and the noise) corresponds to an ill posed inverse problem, due to the huge condition number of the matrix in the forward model. The proposed inversion is based on the general Bayesian inference, based on a generalization of the Student-t distribution used as the sparsity enforcing prior for the PC amplitudes vector.

In Chapter (5) we test the results corresponding to the proposed algorithm on synthetic data, for different levels of noise.

In Chapter (6) the proposed algorithm is used for real data. Finally, in Chapter (7) we present the conclusions.

# 3

## Classical Methods

### Contents

---

<b>3.1 Basic example and notations</b>	<b>27</b>
<b>3.2 Fourier Transform</b>	<b>29</b>
<b>3.3 Zero Padding</b>	<b>31</b>
<b>3.4 Direct or Generalized Inversion</b>	<b>32</b>
<b>3.5 Least Squares and regularized LS</b>	<b>37</b>
<b>3.6 <math>L_1</math> regularization and LASSO</b>	<b>43</b>
<b>3.7 Conclusion and why go to Bayesian</b>	<b>48</b>

---

THE main objective of this chapter is to show that the classical methods have some limitations when faced with the biological signals we are concerned. To show this, first we simulate a synthetic signal  $g(t)$  very near to the biological data we may have to process (short signal, i.e. 4 days length, very noisy, i.e. 5dB) and show the limitations of many well known methods.

### 3.1 Basic example and notations

---

For this purpose, first we consider the following forward model:

$$g(t) = g_0(t) + \epsilon(t), \quad (3.1)$$

where

- $g_0(t) = \sum_{j=1}^M f_j e^{-i\frac{2\pi}{p_j}t}$ ;
- $\mathbf{p} = \{p_1, p_2, \dots, p_M\}$  is the known periods vector, called Periodic Components (PC) vector during the manuscript;

- $\mathbf{f} = \{f_1, f_2, \dots, f_M\}$  is the corresponding amplitudes vector, called Periodic Components (PC) amplitudes vector;
- $\epsilon(t)$  represents the errors and noise.

We are assuming that a few components  $\mathbf{p}^*$  of  $\mathbf{p}$  have non-zero amplitudes  $\mathbf{f}^*$ . In order to analyse the effects of the noise, we will apply the methods not only on the original signal  $g_0(t)$ , i.e. the signal without noise, but also on the noisy signal  $g(t)$ . As we are using sampled signal  $g(t)$  for  $t = 0, \dots, (N-1)\Delta t$ , with  $\Delta t = 1h$ . In the following we can write  $\mathbf{g}_0 = \mathbf{H}\mathbf{f}$  and  $\mathbf{g} = \mathbf{g}_0 + \epsilon$  where  $\mathbf{g} = \{g_0, g_1, \dots, g_{N-1}\}$ ,  $\epsilon = \{\epsilon_0, \epsilon_1, \dots, \epsilon_{N-1}\}$ , and  $H_{nm} = e^{-i\frac{2\pi}{p_m}t_n}$ . To summarize:

- $\mathbf{f}$  - original Periodic Component (PC) amplitudes vector;
- $\mathbf{f}^*$  - the non-zero elements of  $\mathbf{f}$ ;
- $\mathbf{g}_0$  - original signal  $\mathbf{g}_0 = \mathbf{H}\mathbf{f}$
- $\mathbf{g}$  - noisy signal  $\mathbf{g} = \mathbf{H}\mathbf{f} + \epsilon$ ;
- $\hat{\mathbf{f}}_0$  - estimated  $\mathbf{f}$  by any method, corresponding to the original signal,  $\mathbf{g}_0$
- $\hat{\mathbf{f}}$  - estimated  $\mathbf{f}$  by any method, corresponding to a noisy input,  $\mathbf{g}$
- $\delta \mathbf{f}^1$  -  $L_1$  relative estimation error  $\delta \mathbf{f}^1 = \frac{\|\mathbf{f} - \hat{\mathbf{f}}\|_1}{\|\mathbf{f}\|_1}$
- $\delta \mathbf{f}^2$  -  $L_2$  relative estimation error  $\delta \mathbf{f}^2 = \frac{\|\mathbf{f} - \hat{\mathbf{f}}\|_2^2}{\|\mathbf{f}\|_2^2}$
- $\hat{\mathbf{g}}$  - reconstructed signal (estimated data)  $\hat{\mathbf{g}} = \mathbf{H}\hat{\mathbf{f}}$
- $\delta \mathbf{g}^1$  -  $L_1$  relative estimation error  $\delta \mathbf{g}^1 = \frac{\|\mathbf{g} - \hat{\mathbf{g}}\|_1}{\|\mathbf{g}\|_1}$
- $\delta \mathbf{g}^2$  -  $L_2$  relative estimation error  $\delta \mathbf{g}^2 = \frac{\|\mathbf{g} - \hat{\mathbf{g}}\|_2^2}{\|\mathbf{g}\|_2^2}$

Some classical methods are presented in this chapter. Section (3.2) presents the results corresponding to the Fourier Transform. Section (3.3) presents the results corresponding to the zero padding technique. In Section (3.4) we discuss the Direct Inversion (DI). Section (3.5) deals with the results corresponding to the Least Squares and Regularized Least Squares techniques. In Section (3.6) we consider the LASSO technique and present the corresponding results.

We consider a synthetic sparse PC amplitudes vector, defined for the circadian domain and for the potential relevant harmonics, [8 : 32], with non zero values corresponding to 1h, 15h and 23h. The corresponding amplitudes are 0.7, 0.5 and 1. The PC amplitudes vector, the forward operator, the corresponding signal and noisy signal (SNR=5dB) are presented in Figure (3.1). We apply the methods mentioned above both for the signal without noise (Figure (3.1c)) and for the noisy signal, 5dB (Figure (3.1d)). For each PC amplitudes vector estimation, we present the comparison with the theoretical PC amplitudes vector, Figure (3.1a) and also the comparison between the reconstructed signal and the signal without noise. For certain methods (LASSO) we will also consider the variability of the  $L_1$  and  $L_2$  PC amplitudes vector error

estimation for different noise realisation, in order to establish if the method is stable. We also present results corresponding to the behaviour of the  $L_1$  and  $L_2$  PC amplitudes vector error estimation depending on the level of noise (SNR).

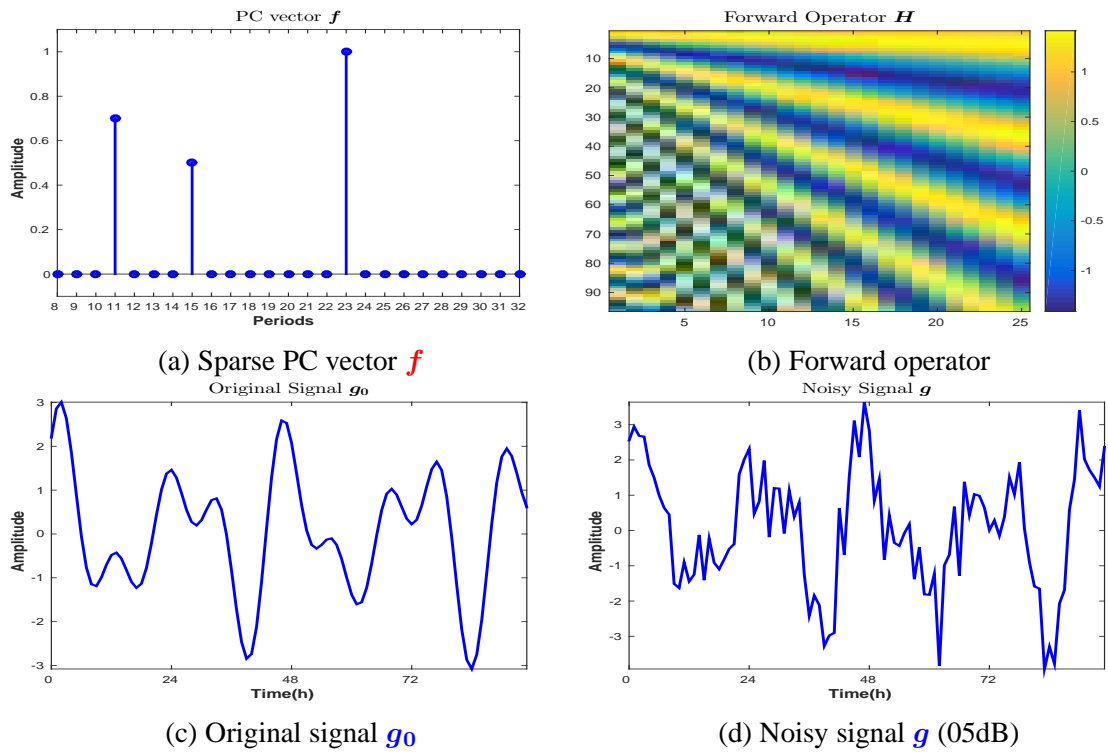


Figure 3.1: Synthetic data: sparse theoretical PC amplitudes vector  $f$  with 3 non-zero peaks corresponding to 11h, 15h and 23h (3.1a); forward operator  $H$  (3.1b); original signal  $g_0 = Hf$  (3.1c); noisy signal  $g = g_0 + \epsilon$  (SNR=05dB) (3.1d)

## 3.2 Fourier Transform

The spectral analysis for time series is a well known subject in literature for a very long time. Some of the most used methods are the Fast Fourier Transform based methods, which are widely used for many applications in the 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 presents certain limitations. In particular, for short time series relative to the searched (via prior knowledge) periodic components (in the experiment considered in this article, a 96h recorded signal relative to a 24h periodic component, linked with the circadian clock) the precision given by the FFT methods is by far insufficient for determining the exact periodic components. This is a consequence of the fact that via the FFT based methods, even if the frequencies are linear, the periods are not equidistant. In particular, for a four day (96h) recorded biomedical signal, beside the 24h corresponding periodic components, the nearest amplitudes in the periodic components vector correspond to the 32h, respectively 19h periodic components. More general, if the prior knowledge sets the principal period around a value  $P$ , it is easy to see that in order to obtain a period 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, i.e.  $(P - 1)P(P + 1)$ . For the chronobiology applications, where the circadian period is around

24h, this is translated as the need to record a signal for 575 days in order to obtain a periodic component vector that contains the 23, 24 and 25 periods. Evidently, the signal should be observed for an exponential bigger interval of time in order to obtain a periodic component vector that contains all the periods from the circadian domain is searched.

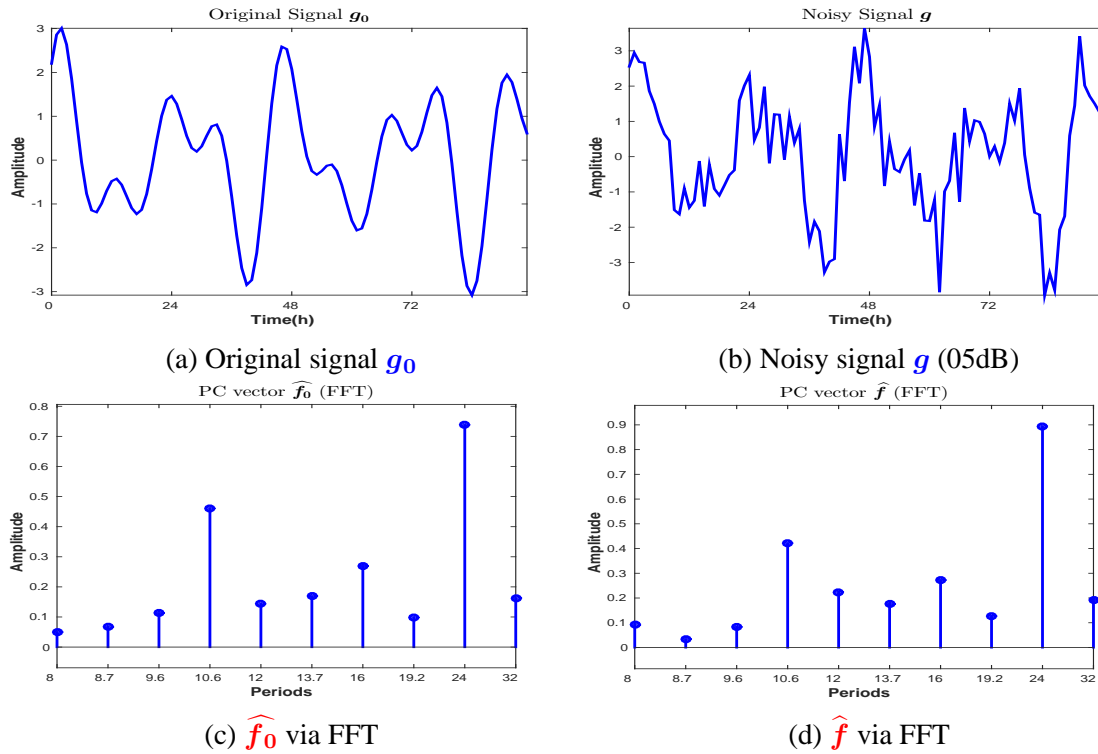


Figure 3.2: Estimated PC amplitudes vectors  $\hat{f}_0$  (3.2c) and  $\hat{f}$  (3.2d) for the original signal  $g$  (3.2a) and the noisy signal  $g_0$  (3.2b) corresponding to the theoretical sparse PC amplitudes vector, (3.1a)

As an example, Figure (3.2) presents the FFT PC amplitudes vectors  $\hat{f}_0$  (Figure (3.2c)) and  $\hat{f}$  (Figure (3.2d)) corresponding to the original signal  $g_0$  and the noisy signal  $g$ . The signals presented in Figure (3.2a) and Figure (3.2b) correspond to the sparse PC vector, Figure (3.1a), via the forward operator, Figure (3.1b).

The corresponding PC amplitudes vectors obtained via FFT are presented for the interval between 8h and 32 h, the circadian domain. In the synthetic PC vector the non-zero periods are set for 11h, 15h and 23h (dominant peak). The FFT estimates the dominant peak at 24h, due to the fact that the time series observation period is limited to 96h, for both cases, original and noisy signal. In such conditions, it offers no information for the correct positions, 11h, 15h, 23h. It also offers no informations for the periods in the interval [20 : 31], except the estimation for 24h. A better image of the intervals with no information from the PC amplitudes vectors estimations is presented in Figure (3.3), where the periods of the PC amplitudes vector are not scaled.

For similar signals, corresponding to PC vectors having the principal peak around 24h, the FFT will estimate the principal peak at 24h. Another example is presented in Figure (3.4). A four days length signal recorded in an experiment in chronobiology is presented in Figure (3.4a). Figure (3.4b) presents its corresponding FFT PC amplitude vector, with the same drawbacks presented in the synthetic example, Figure (3.2). The FFT PC vector presents peaks corresponding only to 8h, 8.72h, 9.6h, 10.66h, 12h, 13.71h, 16h, 19.2h, 24h and 32h

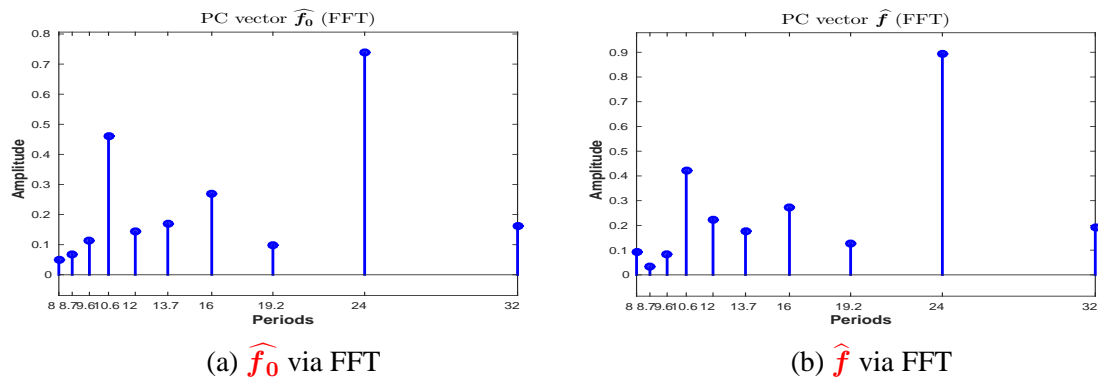


Figure 3.3: Estimated PC amplitudes vectors  $\hat{f}_0$  (3.3a) and  $\hat{f}$  (3.3b). Periods are not scaled in the plot.

inside the considered interval  $[8 : 32]$ . The periods corresponding to 24h, 12h and 8h 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.

Figure (3.4c) presents the PC amplitudes vector estimated via the proposed method, presented in Chapter (4).

### 3.3 Zero Padding

One possible approach to artificially increase the precision of the FFT estimation is the zero padding technique, which consists of appending zeros to the signal. In the simulations presented in this Section (3.2), we have considered a four days length signal, and analysed the FFT PC amplitudes vector estimation, Figure (3.2). Between 8h and 32h, via FFT, we obtained informations corresponding only for 10 periods. Considering the zero padding technique, the number of the periods in the PC amplitudes vector can be increased.

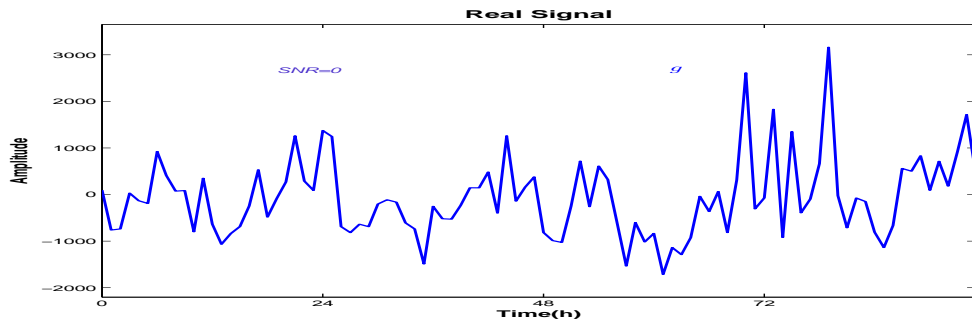
First, we consider the signal from Figure (3.1c) and the noisy signal from Figure (3.1d), both padded with zero values for four days, and the corresponding FFT PC amplitudes vectors, Figure (3.5):

Padding the signal with 4 additional days with zero values, the results do not improve: there are still no informations for certain periods inside the studied domain: compared with the FFT PC amplitudes vector estimated using just the available signal, with zero padding the interval  $[20 : 31]$  has only two additional periods, corresponding to 21.3 and 27.4. A better image of the segments with no information for the periods is showed in Figure (3.6). However, the sparsity rate is increased (the PC amplitudes vector is less sparse), and selecting the peaks corresponding to the biological phenomena is difficult.

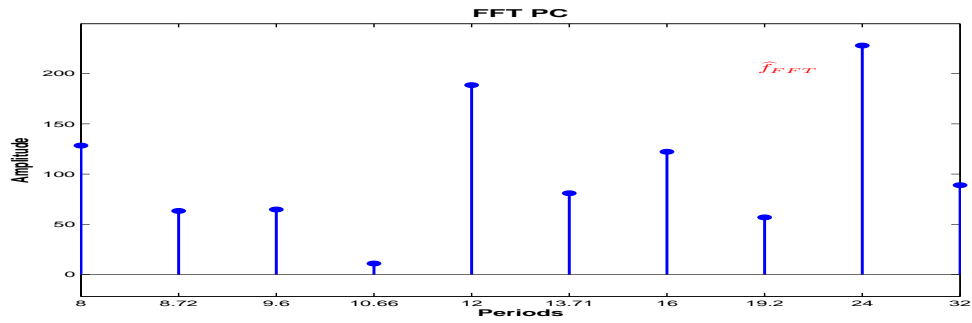
We consider the same approach, padding the signal with 12 days of zero values, Figure (3.7):

Also in this case there are still no informations for certain periods inside the studied do-

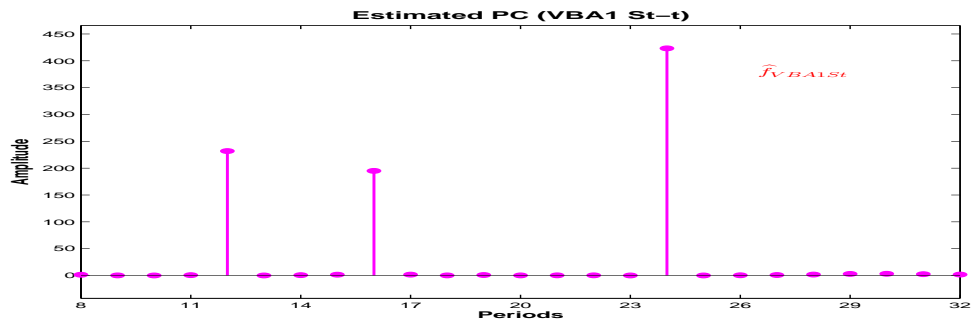




(a) Real signal recorded in cancer treatment experiments,  $f$



(b) FFT PC amplitude vector  $\hat{f}$



(c) PC vector corresponding to Proposed Method (PM estimation, IGSM model)

Figure 3.4: Real data: A 4 days length signal recorded in cancer treatment experiments (3.4a), the PC amplitudes vector corresponding to FFT (3.4b) and corresponding to the proposed method (3.4a).

main, Figure (3.8). Evidently, increasing the length of the zero vector, we can obtain informations corresponding to the all periods from the circadian domain considered. However, the estimation is not a sparse vector, and therefore the selection of the periods corresponding to the biological phenomena is very difficult.

### 3.4 Direct or Generalized Inversion

In this section, we consider another classical method: the Direct Inversion (DI). When  $g = Hf$  and  $H$  is a square and invertible matrix, we have  $\hat{f} = H^{-1}g$ . In fact, the case where  $H$  is the FFT matrix,  $H^{-1}$  is the IFFT matrix. But in our case, the forward operator  $H$  is not necessarily a square matrix, but the pseudoinverse (or the generalized inverse) can still be used. The relation between the original signal  $g_0$  and the corresponding PC amplitudes vector

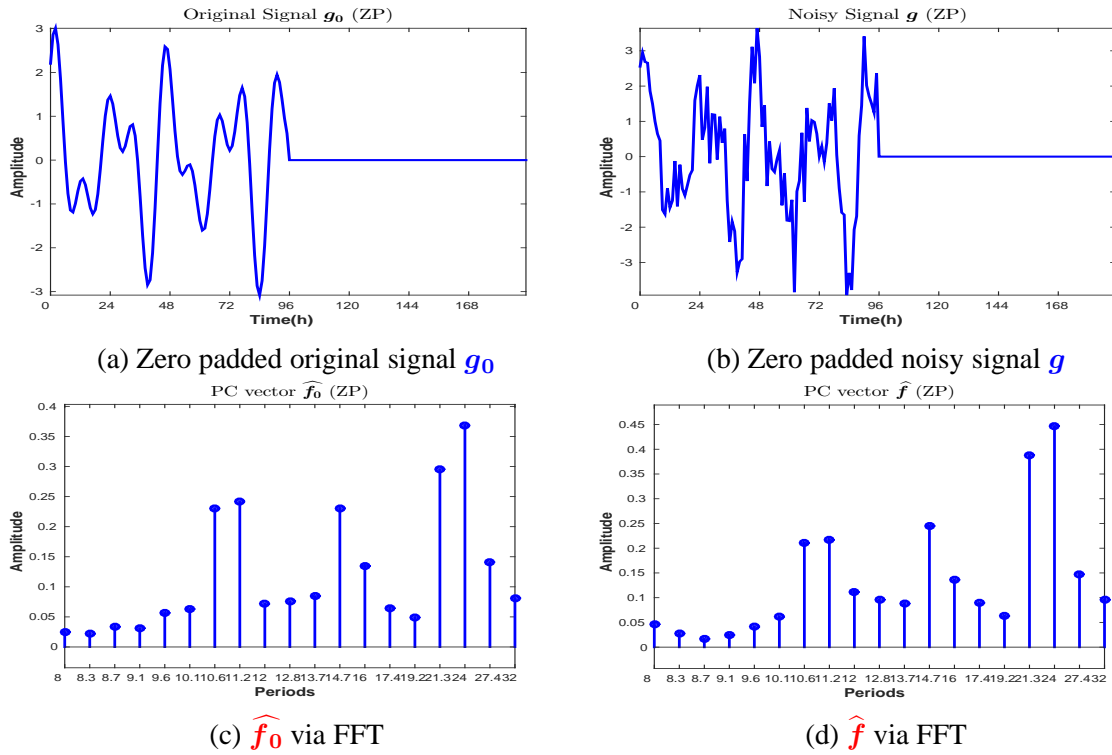


Figure 3.5: Estimated PC amplitudes vectors  $\hat{f}_0$  (3.5c) and  $\hat{f}$  (3.5d) for the zero padded original signal  $g$  (3.5a) and the zero padded noisy signal  $g_0$  (3.5b) corresponding to the theoretical sparse PC amplitudes vector, (3.1a)

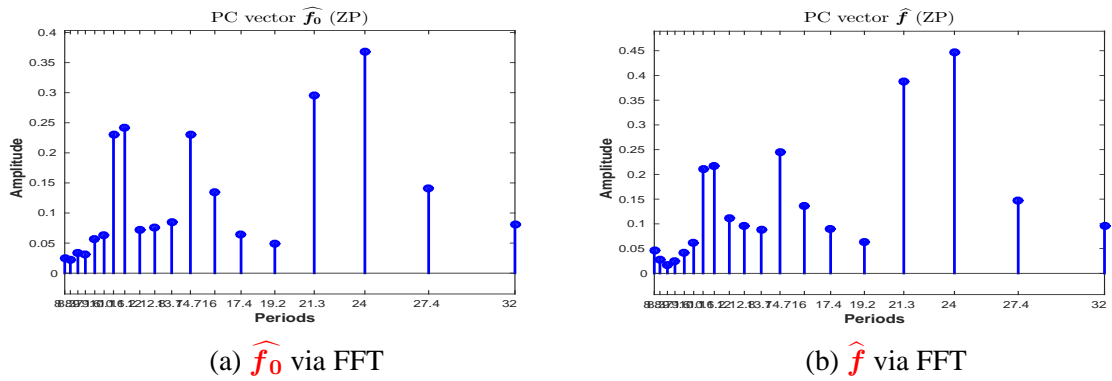


Figure 3.6: Estimated PC amplitudes vectors  $\hat{f}_0$  (3.6a) and  $\hat{f}$  (3.6b), corresponding to zero padded signals. Periods are not scaled in the plot.

$f$  is given in Equation (3.2).

$$g_0 = Hf \tag{3.2}$$

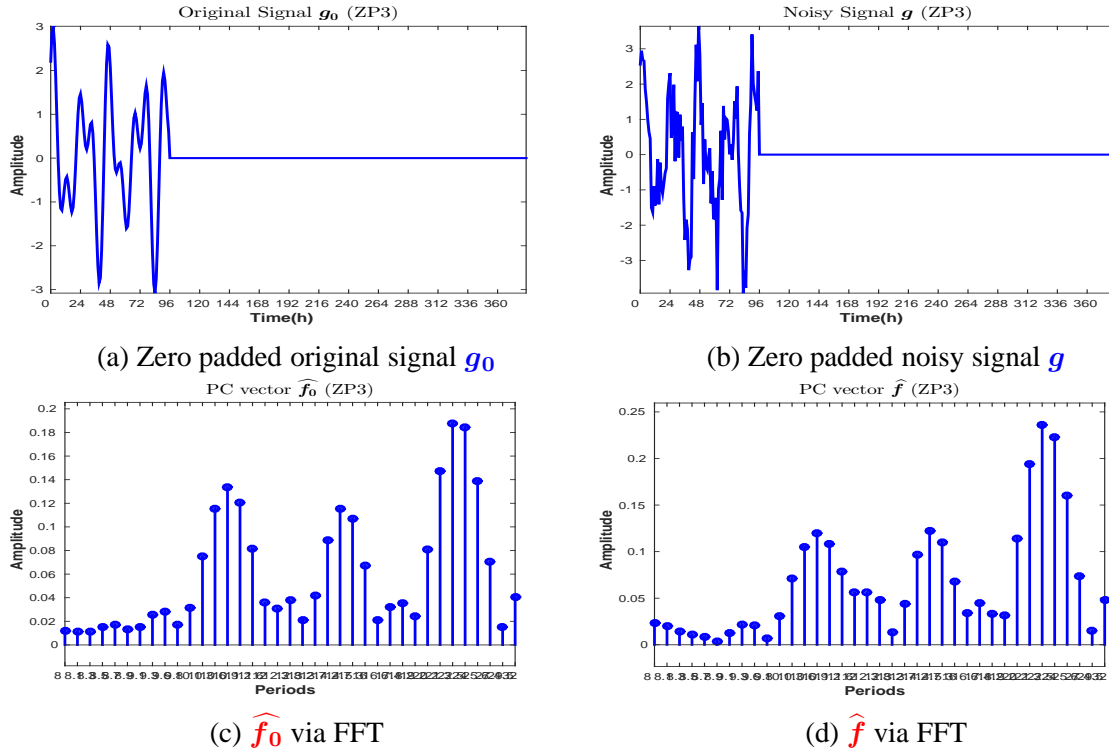


Figure 3.7: Estimated PC amplitudes vectors  $\hat{f}_0$  (3.7c) and  $\hat{f}$  (3.7d) for the zero padded original signal  $g$  (3.7a) and the zero padded noisy signal  $g_0$  (3.7b) corresponding to the theoretical sparse PC amplitudes vector, (3.1a)

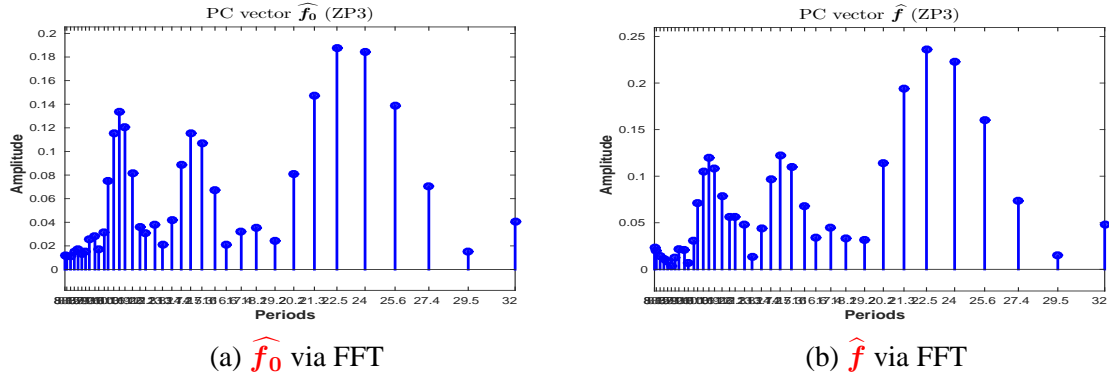


Figure 3.8: Zero padded signal (a) and the corresponding FFT PC amplitudes vector.

Defining  $H^+ = H^T (HH^T)^{-1}$  it is easy to verify that  $H^+$  is satisfying all of the following four criteria:

$$HH^+H = H \quad (3.3a)$$

$$H^+HH^+ = H^+ \quad (3.3b)$$

$$(HH^+)^T = HH^+ \quad (3.3c)$$

$$(H^+H)^T = H^+H \quad (3.3d)$$

so  $H^+$  is the pseudoinverse of  $H$ . First, we apply the method considering the signal without noise, Figure (3.9a). The estimated PC vector via DI is presented in Figure (3.9b).

The comparison between the estimated PC amplitudes vector via DI and the theoretical PC amplitudes vector is presented in Figure (3.9c). The  $L_1$  PC amplitude vector estimation error is  $\delta \mathbf{f}_1 = 0.86 \times 10^{-6} \simeq 0$  and the  $L_2$  PC amplitude vector estimation error is  $\delta \mathbf{f}_2 = 0.56 \times 10^{-6} \simeq 0$ . The comparison between the reconstructed signal the original signal is presented in Figure (3.9d), with the  $L_1$  signal reconstruction error is  $\delta \mathbf{g}_1 = 0.42 \times 10^{-15} \simeq 0$  and the  $L_2$  signal reconstruction error error is  $\delta \mathbf{g}_2 = 0.44 \times 10^{-15} \simeq 0$ . In this case, the DI method gives very good result.

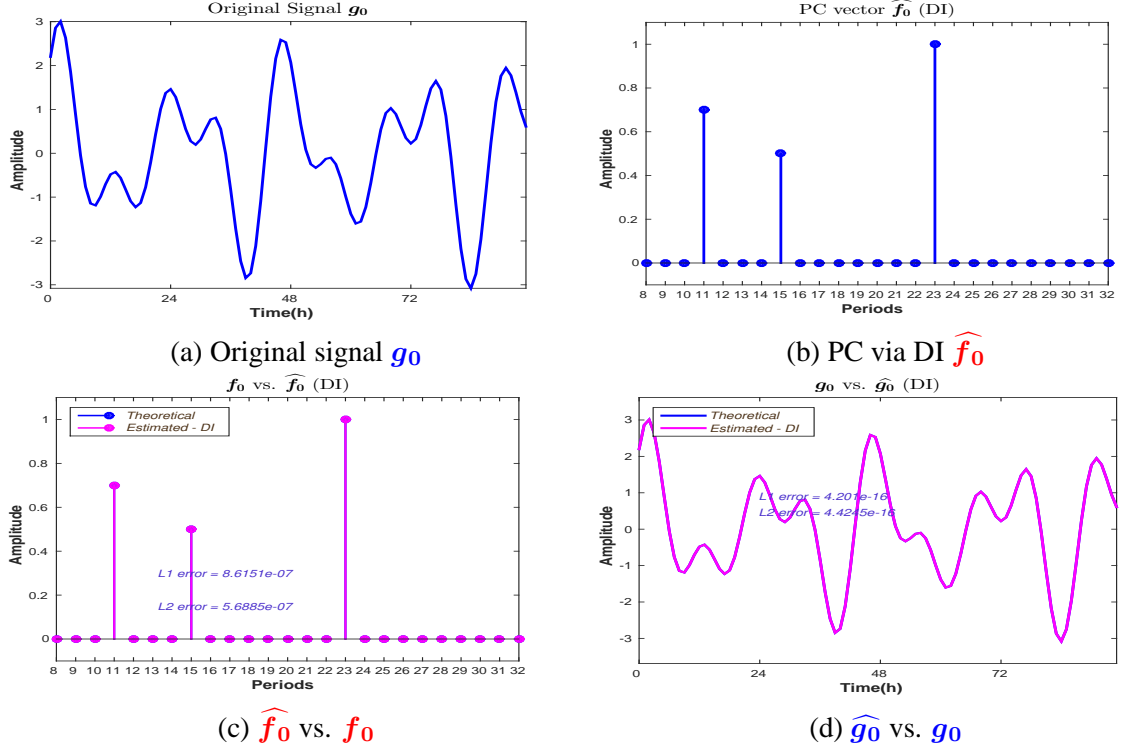


Figure 3.9: Direct Inversion: Signal without noise

However, for the forward operator, we note that the condition number of the matrix is very high. In particular, in this case we have

$$\text{cond}(\mathbf{H}) = 5.6799 \times 10^{10}, \quad (3.4)$$

where  $\mathbf{H}$  denotes the forward operator, Figure (3.1b). Due to this high value of the condition number of the matrix, the problem is ill-posed. We consider in the following the results corresponding to the DI, but applied on the noised signal. First, we consider the case SNR=50dB, Figure (3.10).

We note that for a very small SNR, even though the signal is well reconstructed, Figure (3.10d), with the  $L_1$  signal reconstruction error  $\delta \mathbf{g}_1 = 1.6 \times 10^{-3}$  and the  $L_2$  signal reconstruction error error is  $\delta \mathbf{g}_2 = 1.6 \times 10^{-3}$ , the PC estimation via DI, Figure (3.10c) has a huge error: the  $L_1$  PC amplitude vector estimation error is  $\delta \mathbf{f}_1 = 4.87 \times 10^6$  and the  $L_2$  PC amplitude vector estimation error is  $\delta \mathbf{f}_2 = 3.24 \times 10^6$ . Further, we consider the SNR=20dB (Figure (3.11)) and SNR=05dB (Figure (3.12)) cases:

In both cases, the signal reconstruction is fairly accurate. For 20dB, the  $L_1$  signal reconstruction error  $\delta \mathbf{g}_1 = 3 \times 10^{-2}$  and the  $L_2$  signal reconstruction error error is  $\delta \mathbf{g}_2 = 4 \times 10^{-2}$ , Figure (3.11d) and for 05dB, the  $L_1$  signal reconstruction error  $\delta \mathbf{g}_1 = 2.3 \times 10^{-1}$  and the  $L_2$

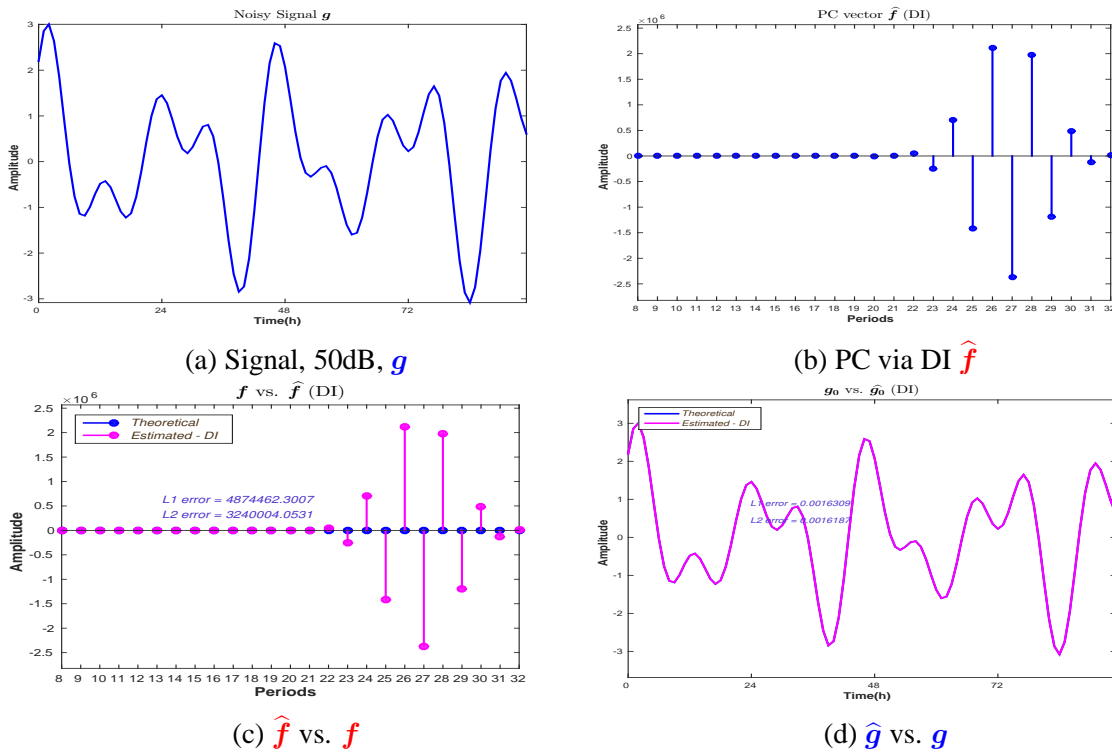


Figure 3.10: Direct Inversion: Signal with noise, 50dB

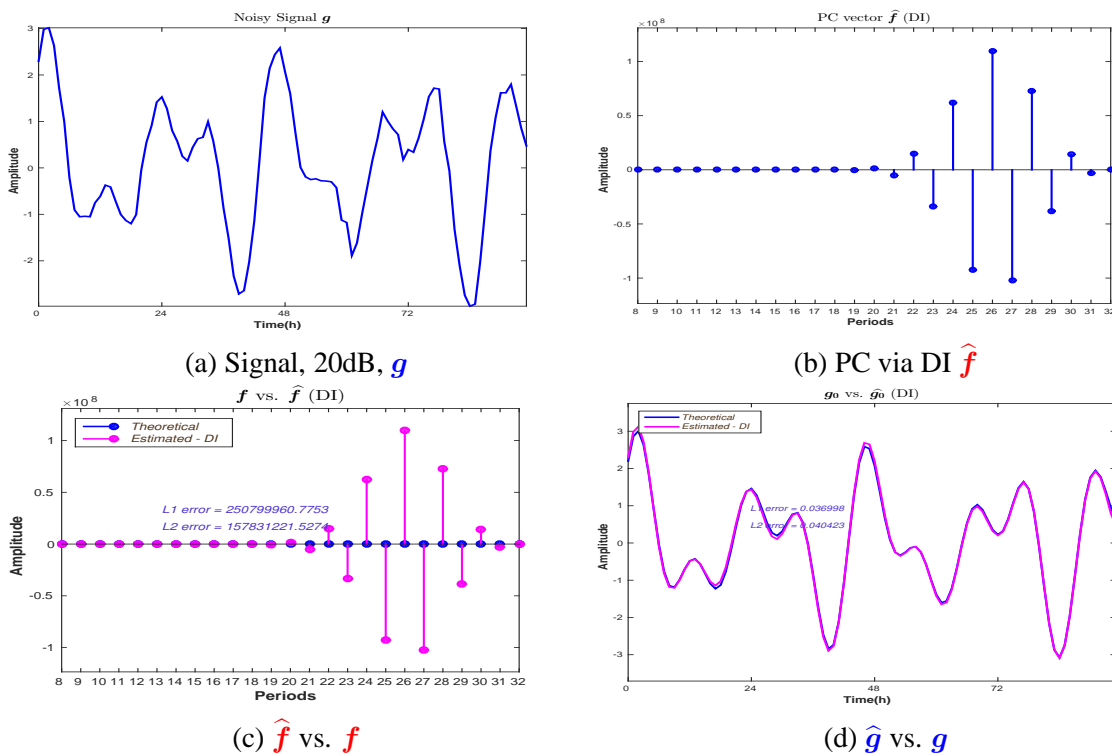


Figure 3.11: Direct Inversion: Signal with noise, 20dB

signal reconstruction error error is  $\delta g_2 = 2.4 \times 10^{-1}$ , Figure (3.12d). In both cases, the PC amplitude vector error estimation is huge. For 20dB the  $L_1$  PC amplitude vector estimation error is  $\delta f_1 = 2.5 \times 10^9$  and the  $L_2$  PC amplitude vector estimation error is  $\delta f_2 = 1.5 \times 10^9$ , Figure (3.11c) and for 05dB the  $L_1$  PC amplitude vector estimation error is  $\delta f_1 = 2.7 \times 10^{10}$

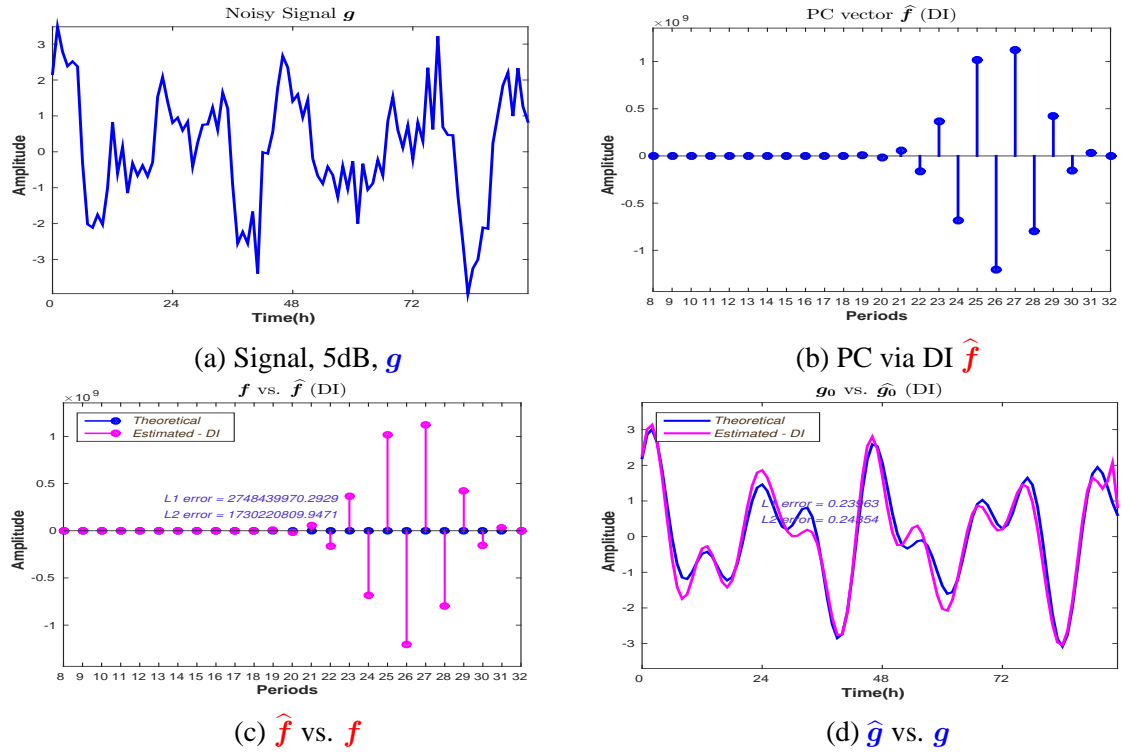


Figure 3.12: Direct Inversion: Signal with noise, 05dB

and the  $L_2$  PC amplitude vector estimation error is  $\delta f_2 = 1.7 \times 10^{10}$ , Figure (3.12c). The behaviour of the DI  $L_1$  and  $L_2$  PC amplitudes vector estimation error  $\delta f_1$  and  $\delta f_2$  and  $L_1$  and  $L_2$  signal reconstruction error  $\delta g_1$  and  $\delta g_2$ , depending on the level of noise is presented in Figure (3.13).

The PC estimation error via Direct Inversion is huge even for small values corresponding to the SNR, Figure (3.13a). Therefore, the Direct Inversion approach can not be used for a precise PC amplitudes vector estimation, due to the fact that the for the real data, the level of noise is very important.

### 3.5 Least Squares and regularized LS

In this section, we consider another classical method: the Least Squares and its extension, regularized Least Squares. For the LS, the solution is given by:

$$\hat{f} = \arg \min_f \|g - Hf\|_2^2, \quad (3.5)$$

which leads to

$$\hat{f} = (H^T H)^{-1} H^T f \quad (3.6)$$

For the regularized LS, the solution is given by:

$$\hat{f} = \arg \min_f J(f), \quad (3.7)$$

where the criterion  $J(f)$  is defined as:

$$J(f) = \|g - Hf\|_2^2 + \lambda \|f\|_2^2. \quad (3.8)$$

## CHAPTER 3. CLASSICAL METHODS

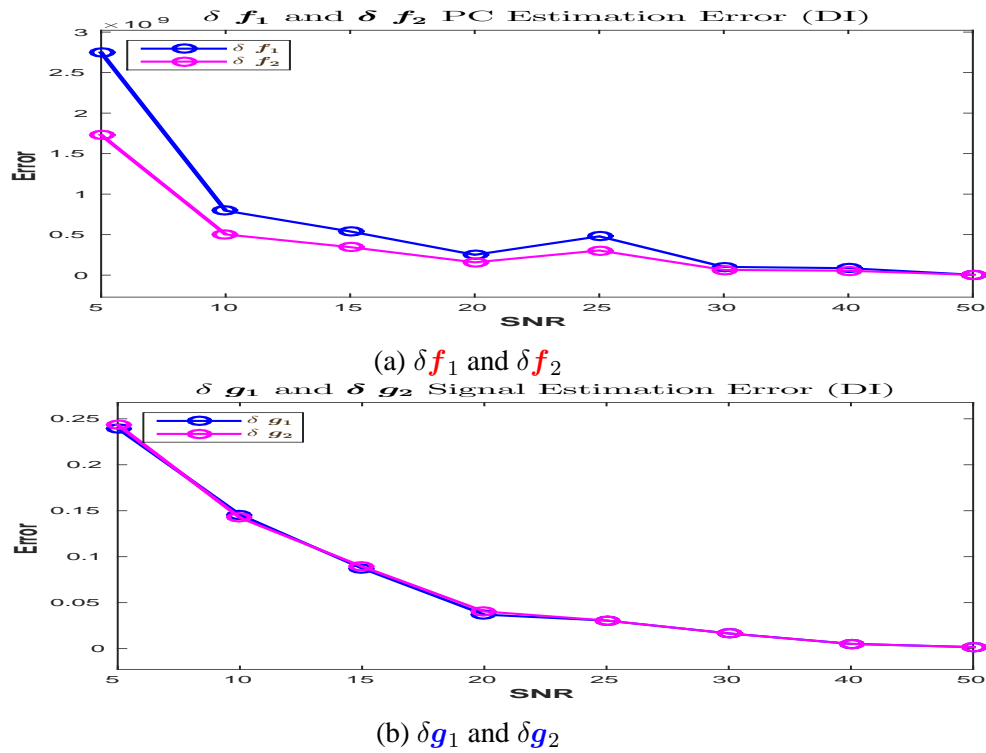


Figure 3.13: Direct Inversion  $L_1$  and  $L_2$  errors: PC estimation errors and signal reconstruction errors

First, we consider the Least Squares, applied on the signal without noise, Figure (3.14).

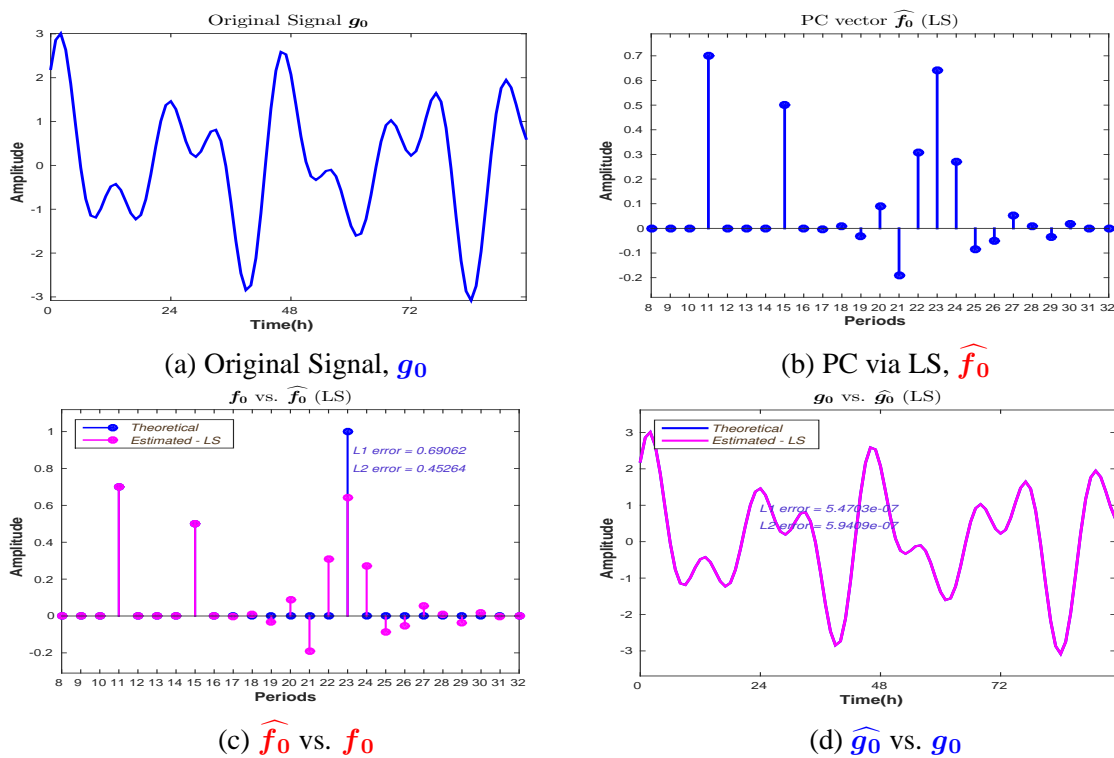


Figure 3.14: Least Squares: Signal without noise

We note that the reconstruction of the signal is accurate, with the  $L_1$  signal reconstruction

error  $\delta g_1 = 5.4 \times 10^{-7} \simeq 0$  and the  $L_2$  signal reconstruction error is  $\delta g_2 = 5.94 \times 10^{-7} \simeq 0$ , Figure (3.14d).

The LS PC amplitude vector estimation, Figure (3.14c) has a significant error: the  $L_1$  PC amplitude vector estimation error is  $\delta f_1 = 0.69$  and the  $L_2$  PC amplitude vector estimation error is  $\delta f_2 = 0.45$ . We note that different from the DI approach, presented in Section (3.4), for which the PC amplitude vector estimation error is huge when the considered signal is noisy, but the PC amplitude vector estimation error is very small for the case when the input has no noise, for the LS approach, the PC amplitude vector estimation error is significant even in this case.

We consider the cases where the signal is noised, SNR= 15dB, Figure (3.15), SNR= 10dB, Figure (3.16) and SNR= 05dB, Figure (3.17)

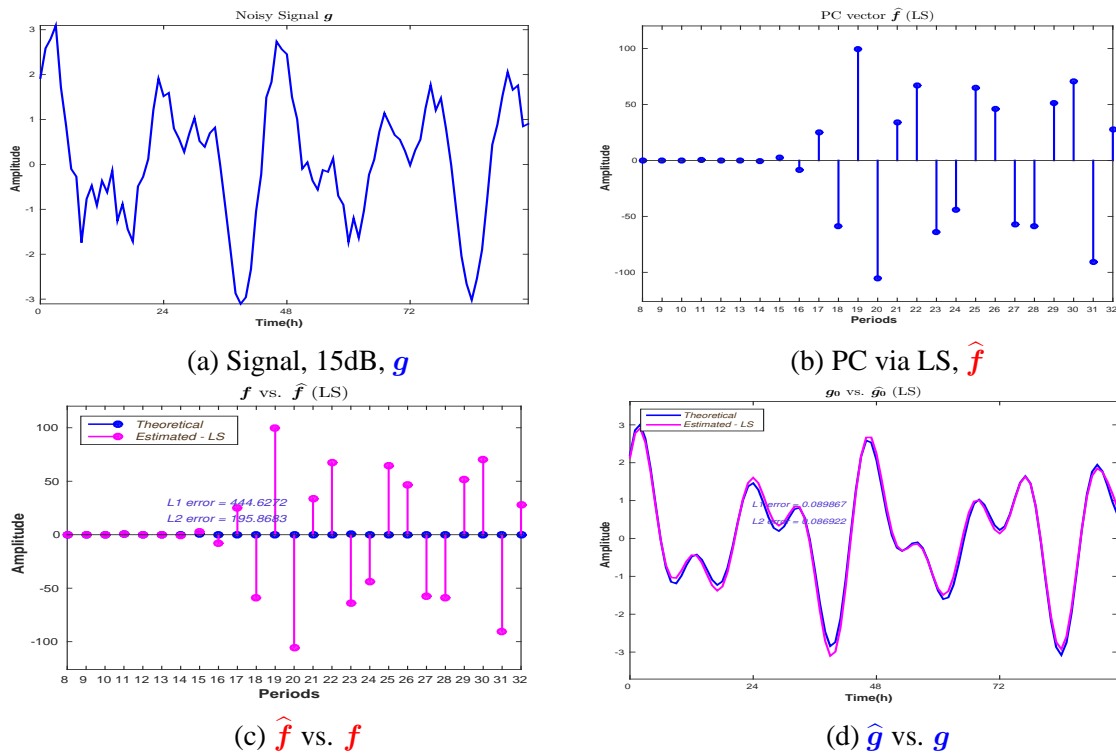


Figure 3.15: Least Squares: Signal with noise, 15dB

The results corresponding to the LS method, considering the input signal with SNR= 15dB noise is presented in Figure (3.15). The input signal, with SNR= 15dB is presented in Figure (3.15a). The estimated PC vector via LS is presented in Figure (3.15b). The comparison between the estimated PC amplitudes vector via LS and the theoretical PC amplitudes vector is presented in Figure (3.15c). The  $L_1$  PC amplitude vector estimation error is  $\delta f_1 = 4.44 \times 10^2$  and the  $L_2$  PC amplitude vector estimation error is  $\delta f_2 = 1.95 \times 10^2$ . The comparison between the reconstructed signal and the original signal is presented in Figure (3.15c), with the  $L_1$  signal reconstruction error  $\delta g_1 = 0.08$  and the  $L_2$  signal reconstruction error  $\delta g_2 = 0.08$ .

The results corresponding to the LS method, considering the input signal with SNR= 10dB noise is presented in Figure (3.16). The input signal, with SNR= 10dB is presented in Figure (3.16a). The estimated PC vector via LS is presented in Figure (3.16b). The comparison between the estimated PC amplitudes vector via LS and the theoretical PC ampli-



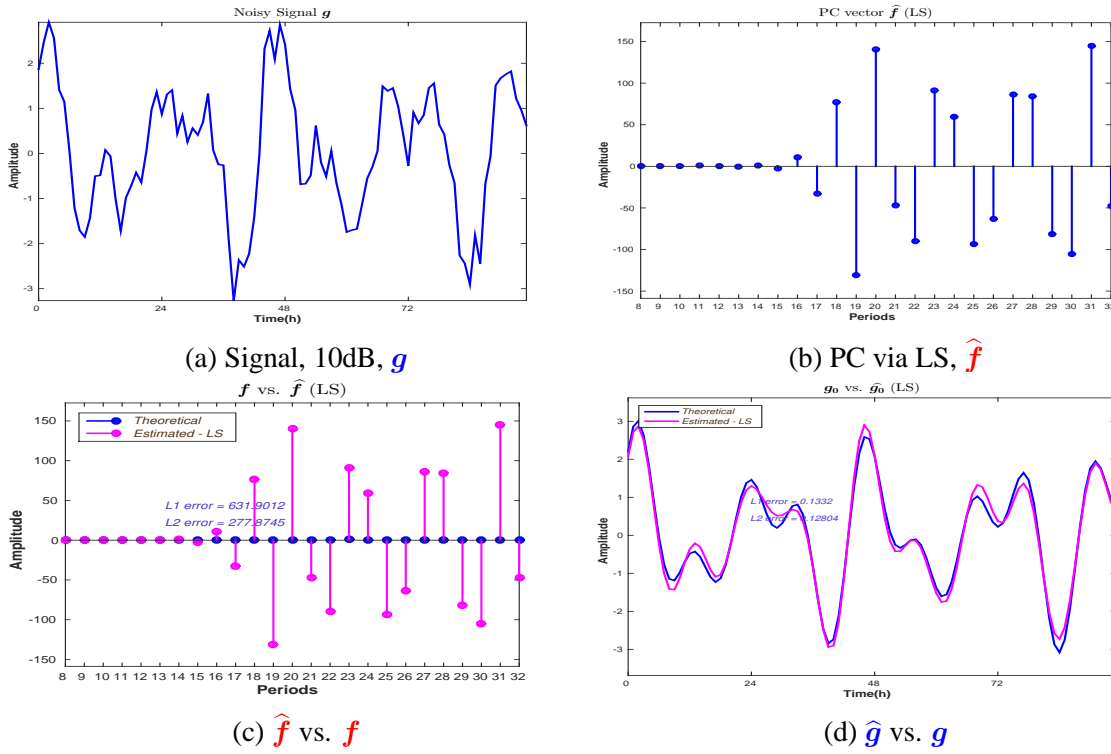


Figure 3.16: Least Squares: Signal with noise, 10dB

tudes vector is presented in Figure (3.16c). The  $L_1$  PC amplitude vector estimation error is  $\delta \mathbf{f}_1 = 6.31 \times 10^2$  and the  $L_2$  PC amplitude vector estimation error is  $\delta \mathbf{f}_2 = 2.77 \times 10^2$ . The comparison between the reconstructed signal and the original signal is presented in Figure (3.16c), with the  $L_1$  signal reconstruction error  $\delta \mathbf{g}_1 = 0.13$  and the  $L_2$  signal reconstruction error  $\delta \mathbf{g}_2 = 0.12$ .

The results corresponding to the LS method, considering the input signal with SNR= 05dB noise is presented in Figure (3.17). The input signal, with SNR= 05dB is presented in Figure (3.17a). The estimated PC vector via LS is presented in Figure (3.17b). The comparison between the estimated PC amplitudes vector via LS and the theoretical PC amplitudes vector is presented in Figure (3.17c). The  $L_1$  PC amplitude vector estimation error is  $\delta \mathbf{f}_1 = 7.16 \times 10^2$  and the  $L_2$  PC amplitude vector estimation error is  $\delta \mathbf{f}_2 = 3.17 \times 10^2$ . The comparison between the reconstructed signal and the original signal is presented in Figure (3.17c), with the  $L_1$  signal reconstruction error  $\delta \mathbf{g}_1 = 0.26$  and the  $L_2$  signal reconstruction error  $\delta \mathbf{g}_2 = 0.25$ .

For all three levels of noise considered, the signal reconstruction errors are fairly accurate. For the PC amplitudes vector estimation errors are very important. The behaviour of the LS  $L_1$  and  $L_2$  PC amplitudes vector estimation error  $\delta \mathbf{f}_1$  and  $\delta \mathbf{f}_2$  and  $L_1$  and  $L_2$  signal reconstruction error  $\delta \mathbf{g}_1$  and  $\delta \mathbf{g}_2$ , depending on the level of noise is presented in Figure (3.18).

An extension of the LS approach is the regularized LS. We consider, different values for the regularization term  $\lambda$  and apply the RLS both on the signal without noise and with noise, SNR= 5dB. Figure (3.19) presents the simulations corresponding to  $\lambda = 0.1$ , considering the signal without noise.

The signal is well reconstructed, Figure (3.19d).  $L_1$  and  $L_2$  reconstruction errors are 0.0018 and 0.002. However, for the PC amplitudes vector estimation, both  $L_1$  and  $L_2$  errors

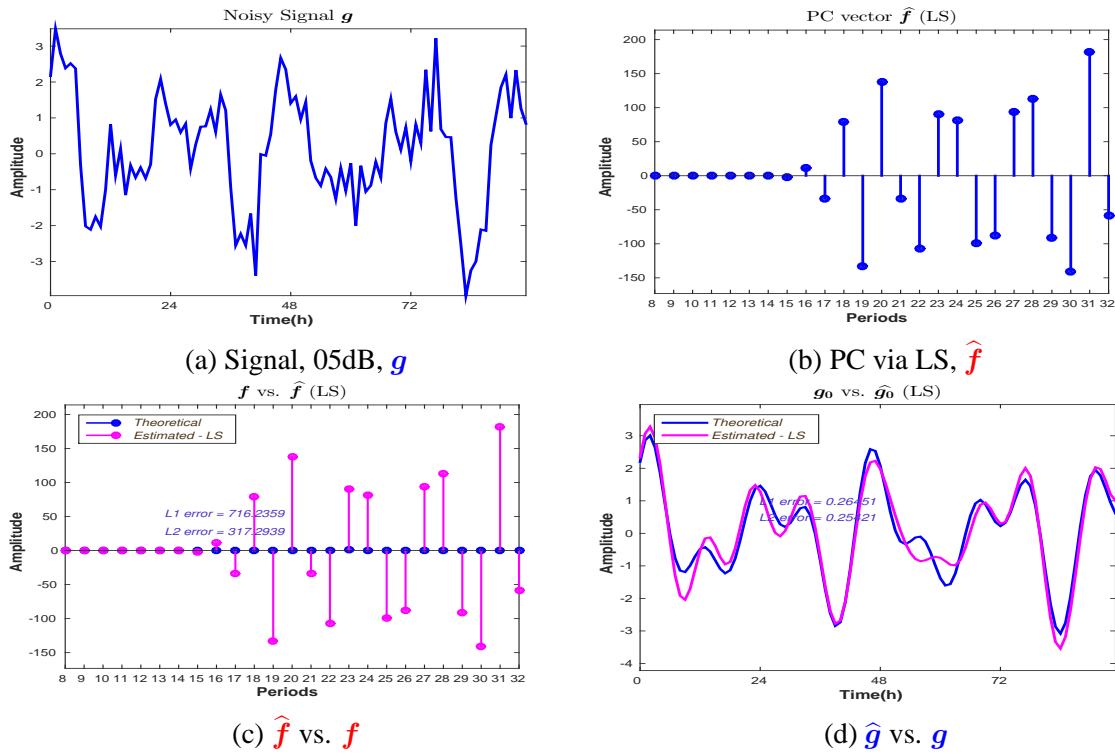


Figure 3.17: Least Squares: Signal with noise, 05dB

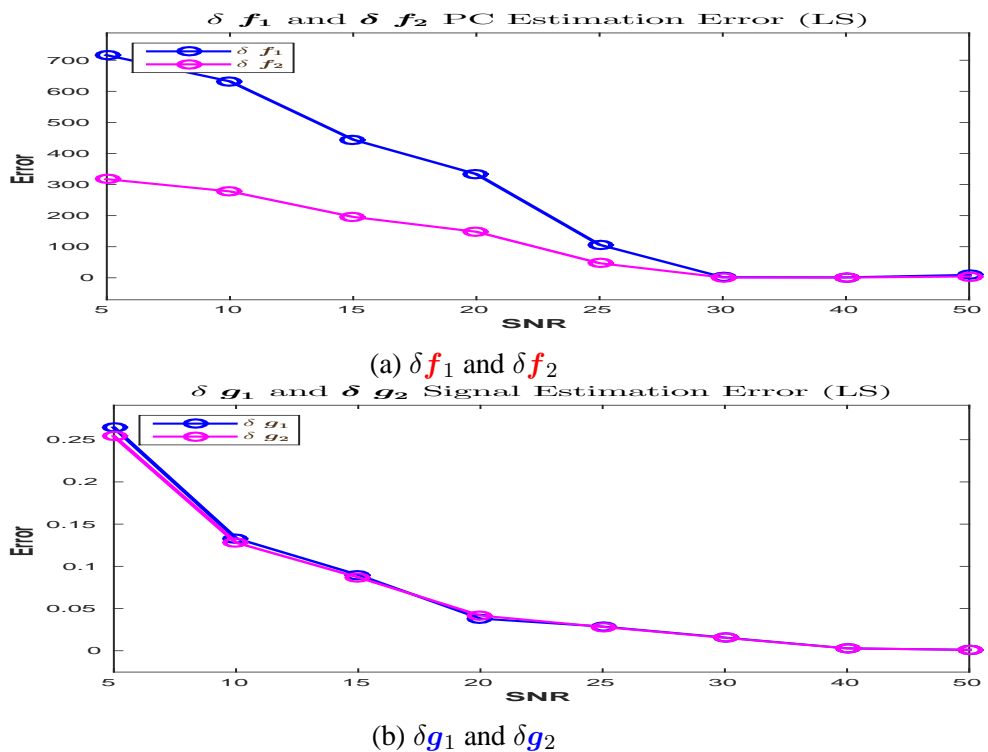


Figure 3.18: Least Squares  $L_1$  and  $L_2$  errors: PC estimation errors and signal reconstruction errors

are high, Figure (3.19c): 0.89 and 0.57.

Figure (3.20) presents the simulations corresponding to the same regularization parame-

## CHAPTER 3. CLASSICAL METHODS

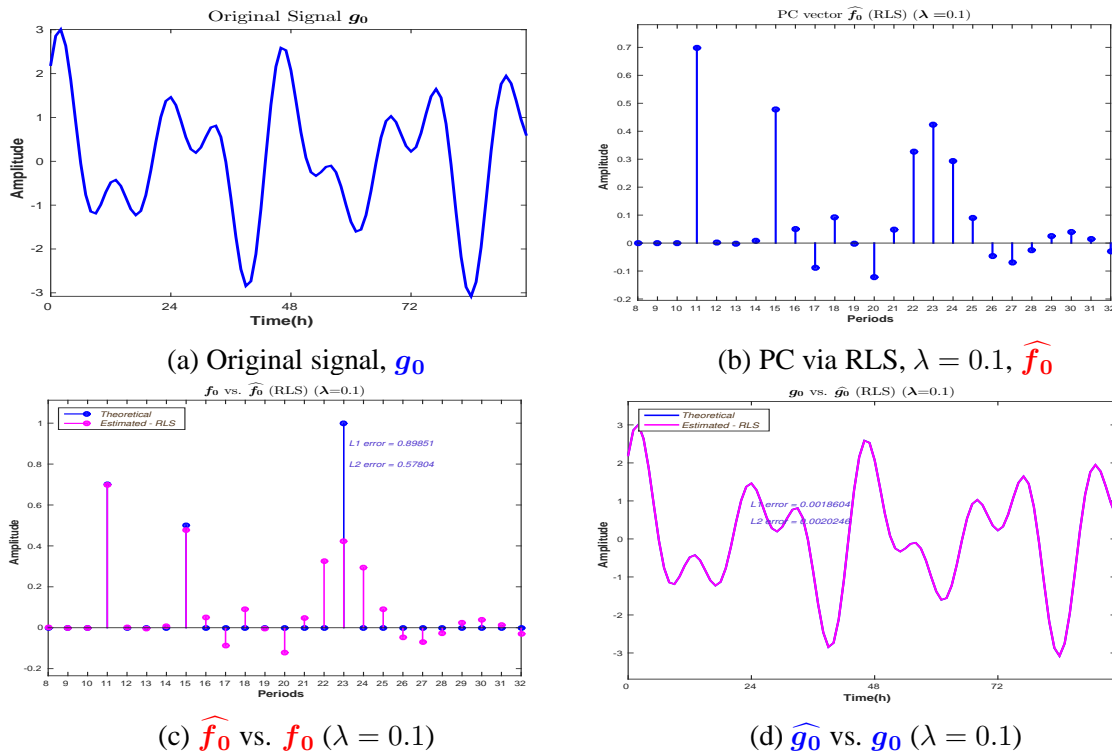


Figure 3.19: Regularized Least Squares,  $\lambda = 0.1$ : Signal without noise

ters, but considering the noisy signal, SNR= 5dB.

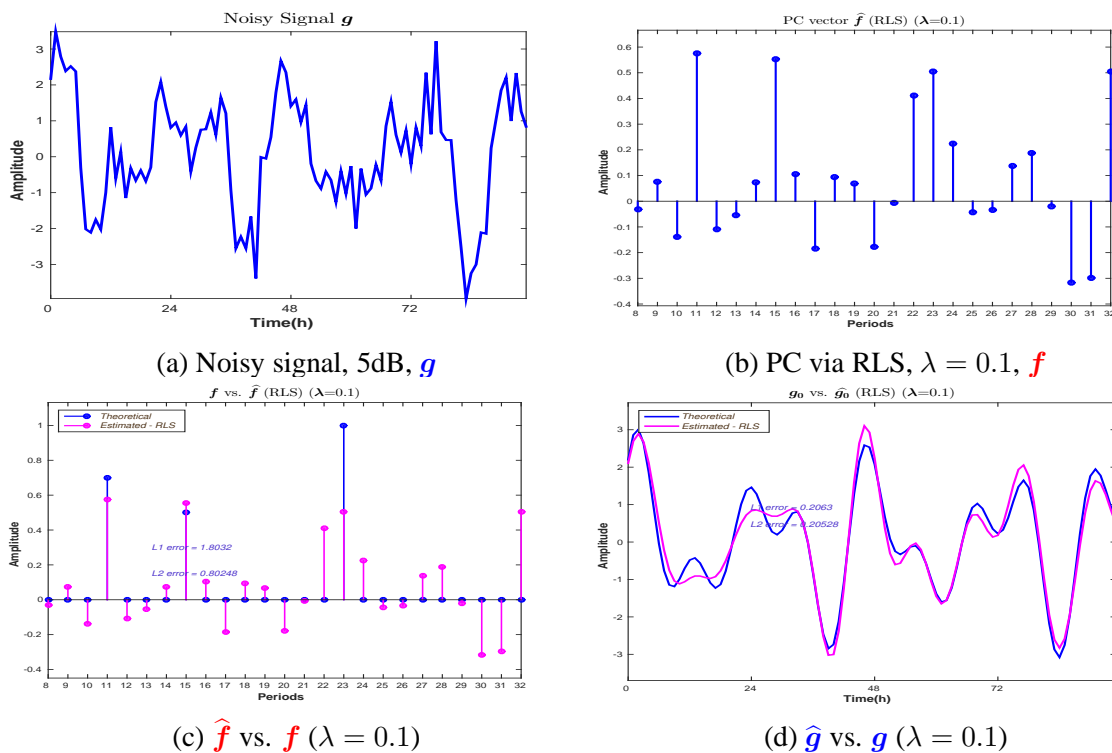


Figure 3.20: Regularized Least Squares,  $\lambda = 0.1$ : Signal with noise, 5dB

The signal is not well reconstructed, Figure (3.20d),  $L_1$  and  $L_2$  reconstruction errors are 0.20 and 0.20, nor the PC amplitudes vector estimation is accurate, both  $L_1$  and  $L_2$  errors are

high, Figure (3.20c): 1.44 and 0.64.

The behaviour of the  $L_1$  and  $L_2$  PC amplitudes vector estimation errors and signal reconstruction are presented in Figure (3.21)

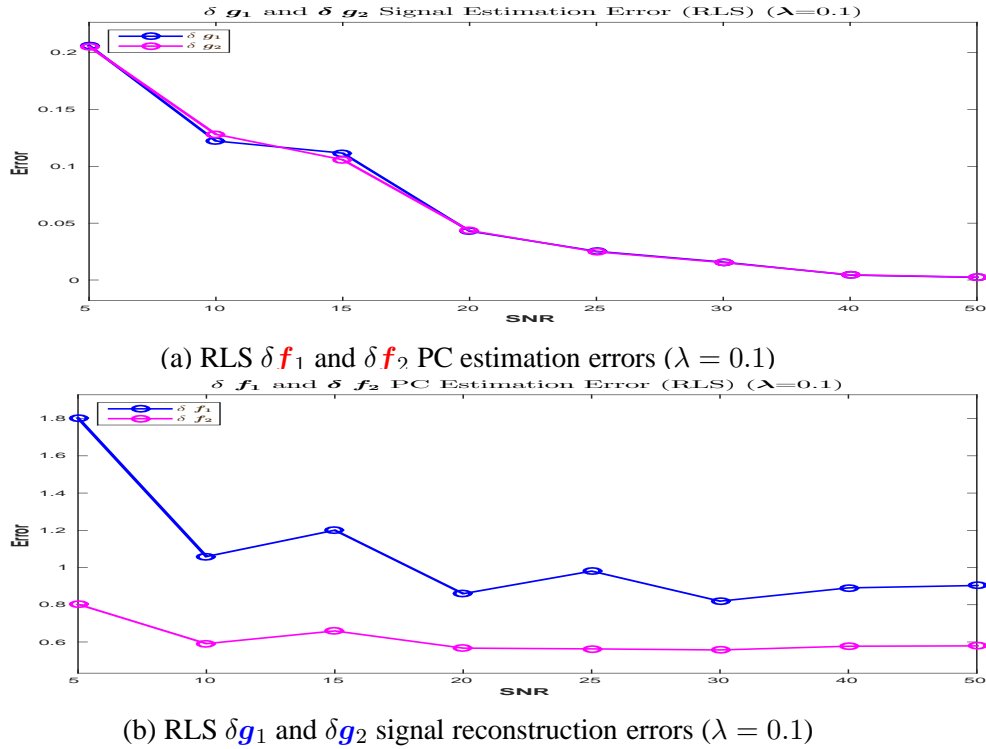


Figure 3.21: Regularized Least Square  $L_1$  and  $L_2$  errors: PC estimation and signal reconstruction

Figure (3.22) presents the behaviour of the L1 (Figure (3.22a)) and L2 (Figure (3.22b)) signal reconstruction error depending on the level of noise, for different values of  $\lambda$  ( $\lambda = \{0.01, 0.1, 0.5, 1, 10\}$ ):

Figure (3.23) presents the behaviour of the L1 (Figure (3.23a)) and L2 (Figure (3.23b)) PC estimation error depending on the level of noise, for different values of  $\lambda$  ( $\lambda = \{0.01, 0.1, 0.5, 1, 10\}$ ):

### 3.6 $L_1$ regularization and LASSO

In Section (3.5), we have considered the LS method, and also the quadratic regularized version of this method. Another regularized version of LS is the Least Absolute Shrinkage and Selection Operator (LASSO) method which uses the constraint that the L1-norm of the PC amplitudes vector is no greater than a given value. We note that in the Bayesian approach, the LASSO method is equivalent to the use of a zero-mean Laplace prior distribution on the PC amplitudes vector. The optimization problem may be solved using quadratic programming or more general convex optimization methods, as well as by specific algorithms such as the LASSO algorithm. The L1-regularized formulation is useful in the context of a sparse PC amplitudes vector, due to its tendency to prefer solutions with a small number of non-zero parameter values.

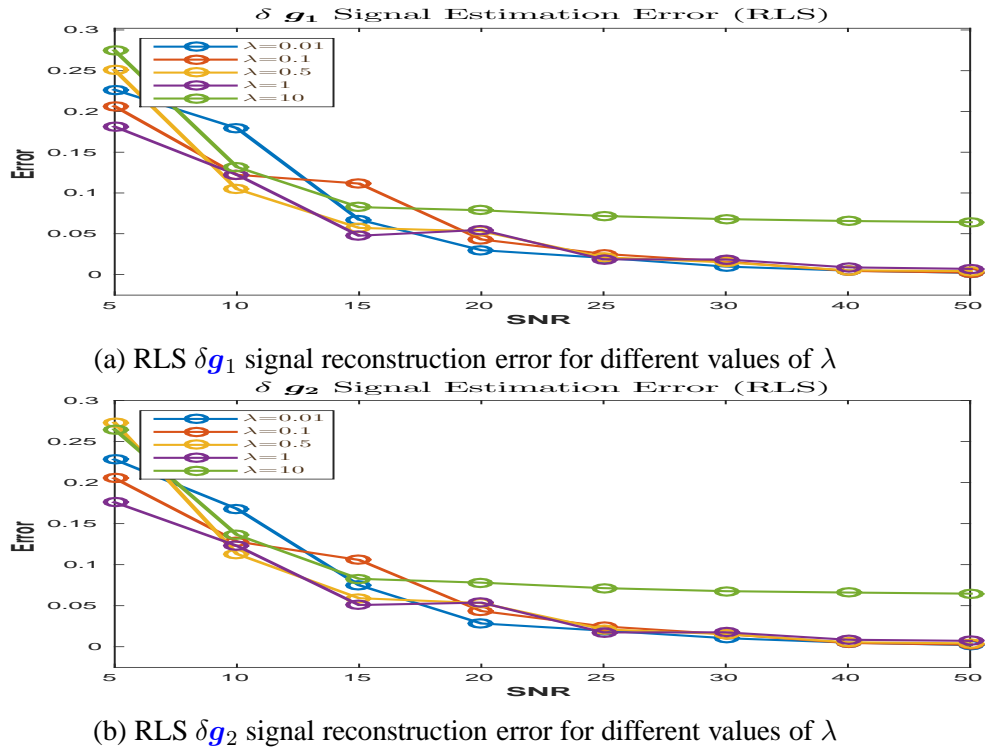


Figure 3.22: Regularized Least Square  $\delta g_1$  and  $\delta g_2$  signal reconstruction errors for different values of  $\lambda$

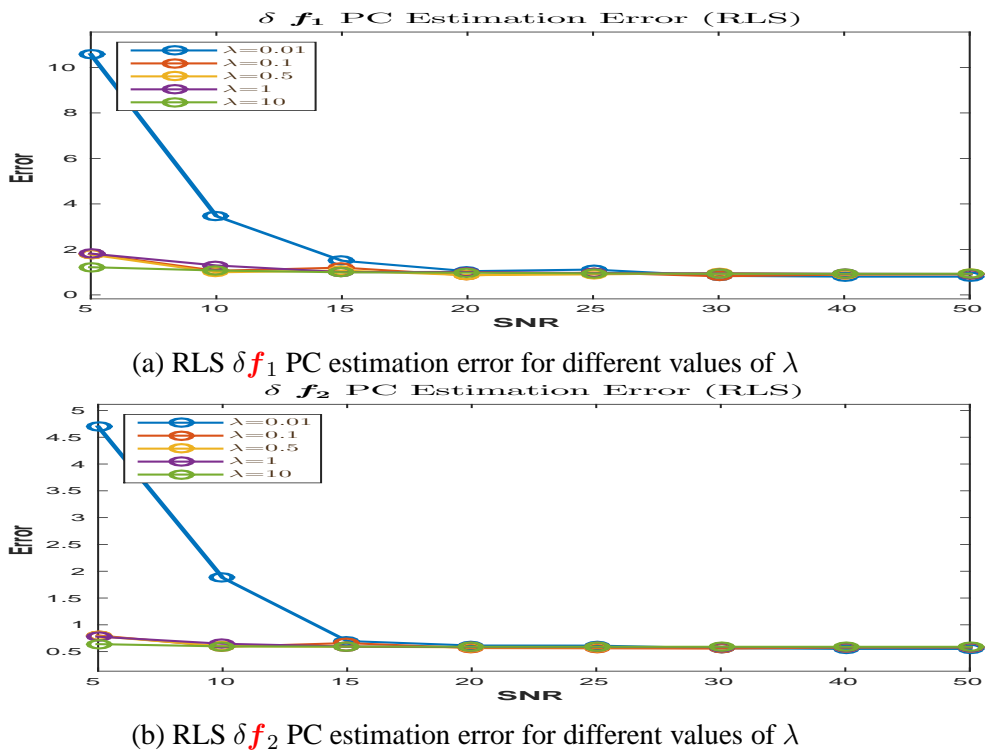


Figure 3.23: Regularized Least Square  $\delta f_1$  and  $\delta f_2$  PC estimation errors for different values of  $\lambda$

For LASSO, the solution is given by:

$$\hat{f} = \arg \min_{f} J(f), \tag{3.9}$$

where the criterion  $J(\mathbf{f})$  is defined as:

$$J(\mathbf{f}) = \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_2^2 + \lambda\|\mathbf{f}\|_1. \quad (3.10)$$

Figure (3.24) presents the simulations corresponding to LASSO method, considering the signal without noise:

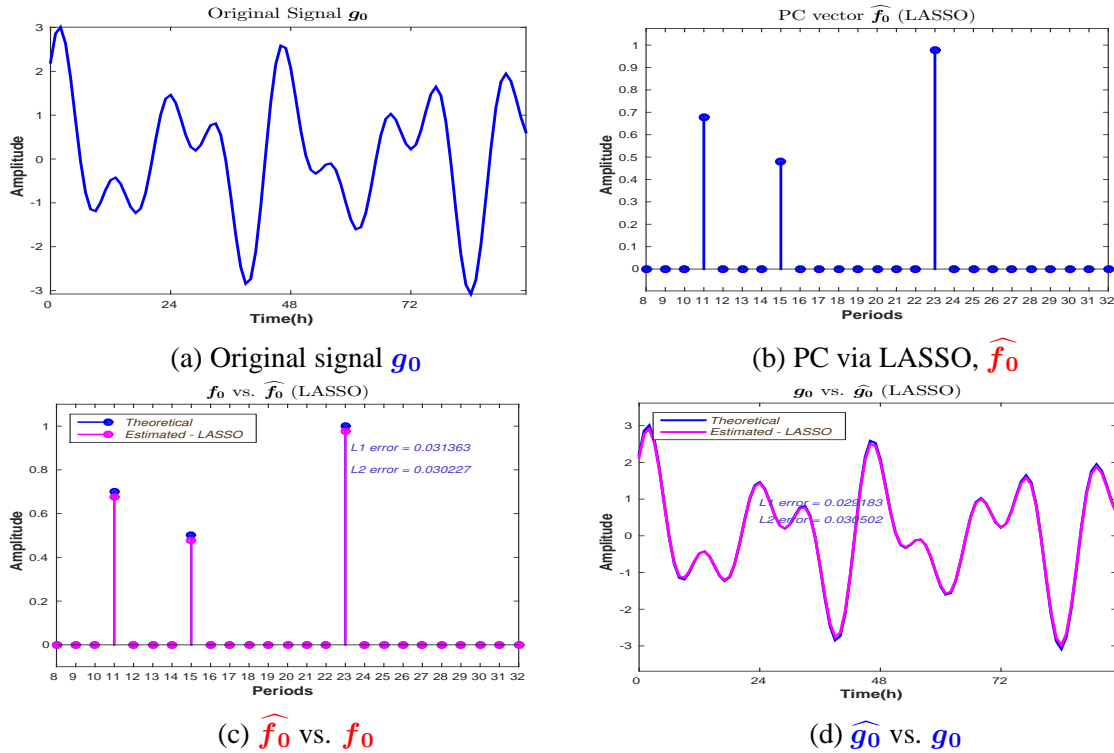


Figure 3.24: LASSO: Signal without noise

In this case, for the signal without noise, the PC amplitudes estimation (and, consequently, the signal reconstruction) is very accurate. Via the LASSO method, the estimation is a sparse vector, and for all the non-zero values, the estimated amplitudes are very close to the real ones, Figure (3.24c). The  $L_1$  error norm is 0.031 and the  $L_2$  error norm is 0.030. The comparison between the reconstructed signal and the original signal is presented in (3.24d). We consider now the case when the input signal is noisy. Figure (3.25) presents the simulations corresponding to LASSO method, but considering the noisy signal, SNR= 10dB.

We note that for a SNR= 10dB, via the LASSO method the estimation of the PC amplitudes vector is still accurate, Figure (3.25c). The estimation is sparse, and except the amplitudes corresponding to 8h and 10h, which are negligible, the periods corresponding to non-zero amplitudes are correctly detected. The  $L_1$  error norm is 0.22 and the  $L_2$  error norm is 0.13. We consider the case when the SNR= 5dB. Figure (3.26) presents the simulations corresponding to LASSO method, but considering the noisy signal with SNR= 5dB.

Figure (3.26c) shows the comparison between the estimated PC amplitudes vector and the theoretical PC amplitudes vector. The  $L_1$  error norm is 0.17 and the  $L_2$  error norm is 0.19. The PC amplitudes vector is sparse and the periods corresponding to non-zero amplitudes are correctly detected, and the amplitudes are fairly accurate. However, LASSO method can provide very inaccurate estimations. Figure (3.27) presents the results corresponding to a simulation done in the same conditions as the one from Figure (3.26): a noisy signal with SNR= 5dB

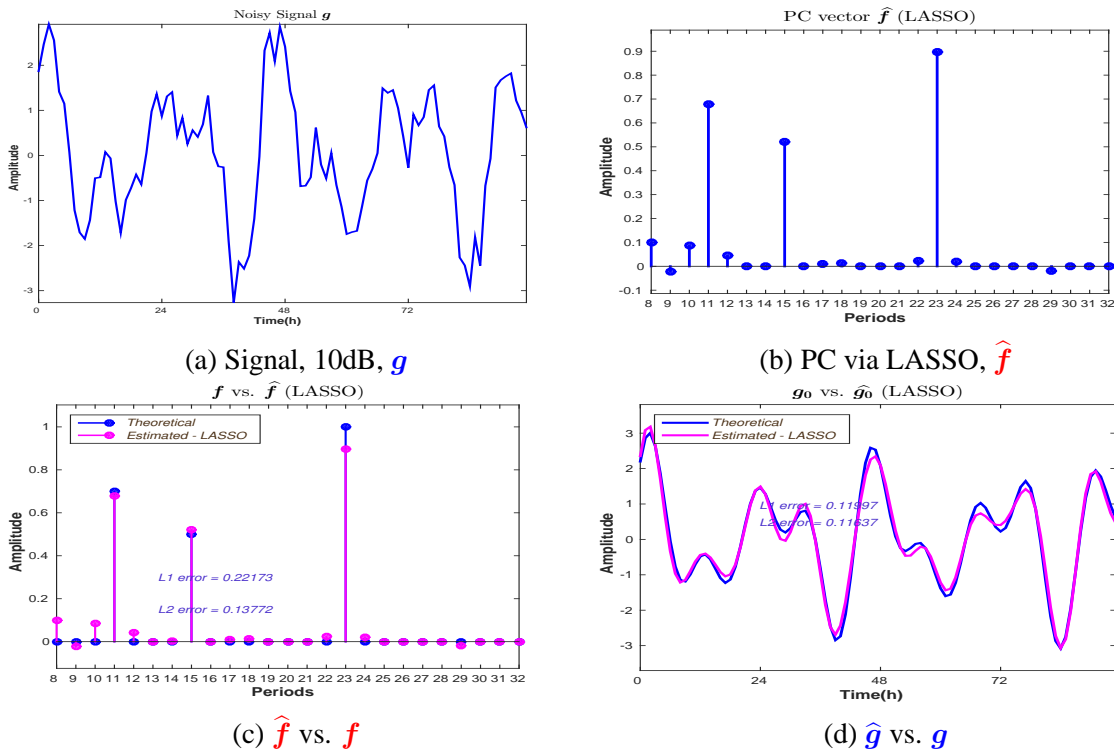


Figure 3.25: LASSO: Signal with noise, 10dB

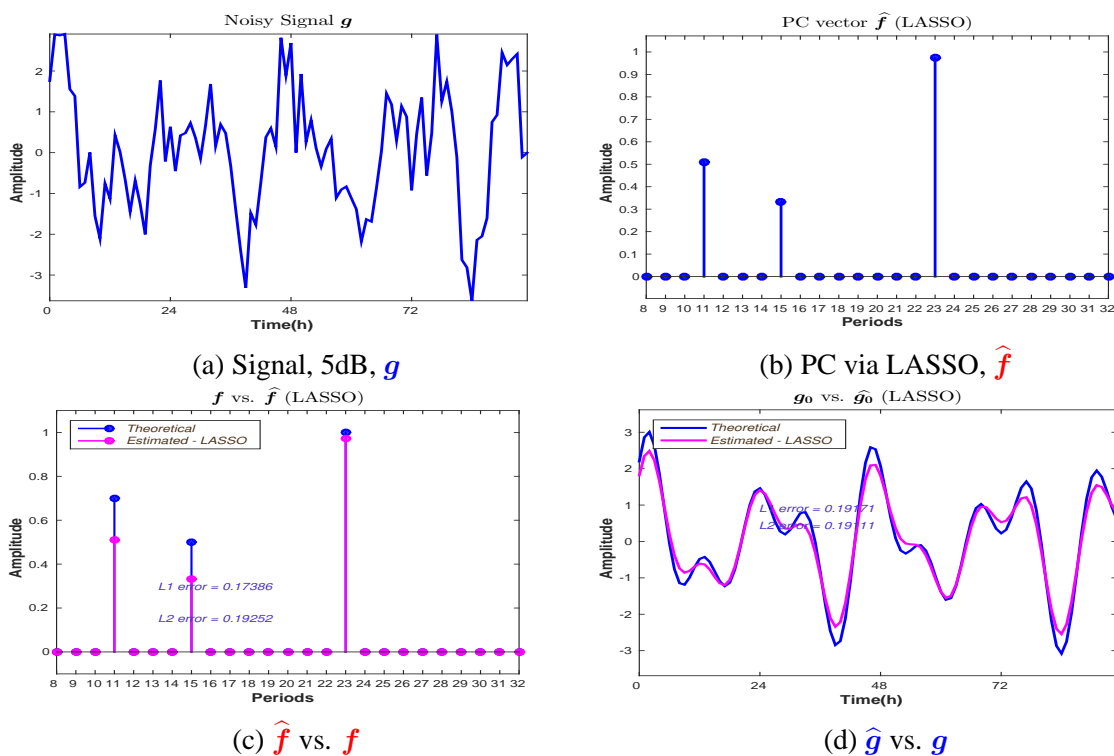


Figure 3.26: LASSO: Signal with noise, 5dB - Accurate estimation

used as the input using the same LASSO method, but for a different noise realisation. However, in this case the solution is not sparse, and the estimation of the amplitude corresponding to the 23h period is far from being accurate, Figure (3.27c). In this case, the  $L_1$  error norm is 3.07 and the  $L_2$  error norm is 1.5.

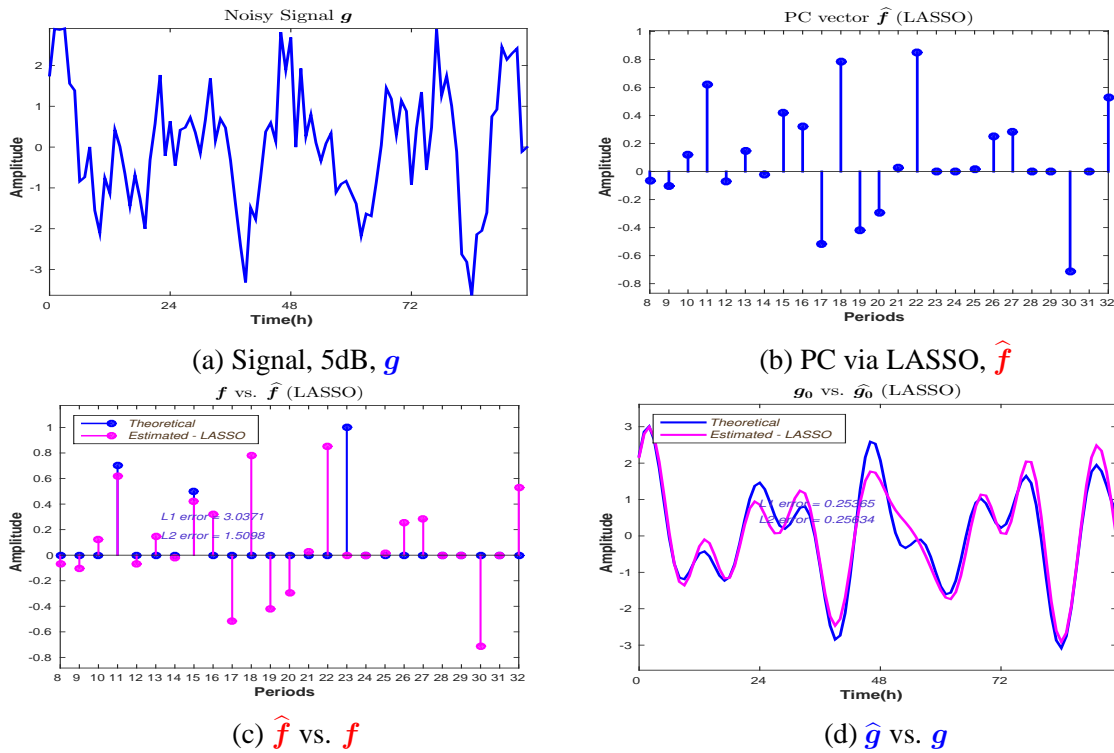


Figure 3.27: LASSO: Signal with noise, 5dB - Inaccurate estimation

LASSO method can provide good results for the PC amplitudes vector, even for SNR= 5dB. But in certain cases, for high signal to noise ratios, the method is not stable. Figure (3.28) presents the variations of the LASSO  $L_1$  and  $L_2$  error norms corresponding to the PC amplitudes vector estimation (Figure (3.28a)) and to the signal reconstruction (Figure (3.28b)) using as an input a noisy signal with SNR= 10dB, for 50 different noise realisations. For the signal reconstruction, the variation is not very important, i.e. the LASSO method provides an estimation of the PC amplitudes vector that can accurately reconstruct the signal. However, for the estimation of the PC amplitudes vector, the variation of the error is important, and in certain cases, the error value is high. In this experiment, we note that for the 50th noise realisation, via the LASSO method the PC amplitudes vector error estimation is superior to 2.5 for  $L_1$  norm and superior to 1.5 for  $L_2$  norm.

The instability of the LASSO method is more obvious when considering the variability of the  $L_1$  and  $L_2$  error norms for 50 different noise realisations with the input the noisy signal, with SNR= 5dB, Figure (3.29) and Figure (3.30).

For the signal reconstruction, in both examples, Figure (3.29b) and Figure (3.30b) the variation is not very important, i.e. the LASSO method provides an estimation of the PC amplitudes vector that can accurately reconstruct the signal. For the estimations of the PC amplitudes vector, the variation of the error is important, and in certain cases, the error value is high. In these experiments, we note that for the 45th noise realisation in the first experiment (Figure (3.29b)) and for the 9th noise realisation in the second experiment (Figure (3.30b)), via the LASSO method the PC amplitudes vector error estimation is superior to 2.5 for  $L_1$  norm for the first experiment and superior to 7 for  $L_1$  norm for the second experiment.

The behaviour of the LASSO  $L_1$  and  $L_2$  error norms corresponding to the PC amplitudes vector estimation and to the signal reconstruction is presented in Figure (3.31).



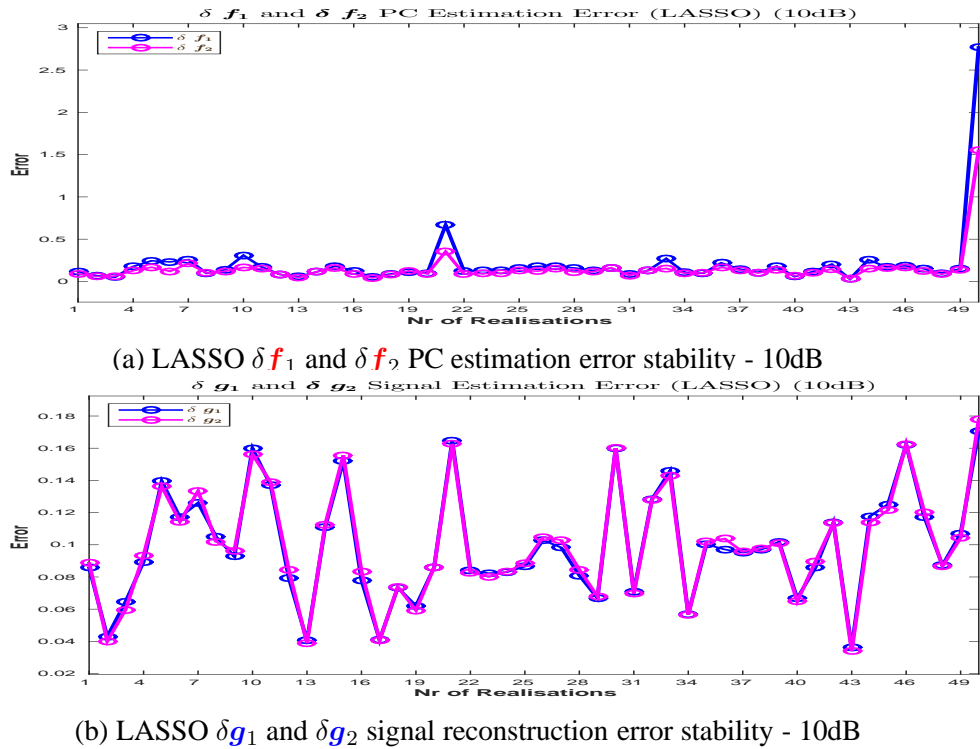


Figure 3.28: LASSO  $L_1$  and  $L_2$  errors stability: PC estimation and signal reconstruction - 10dB

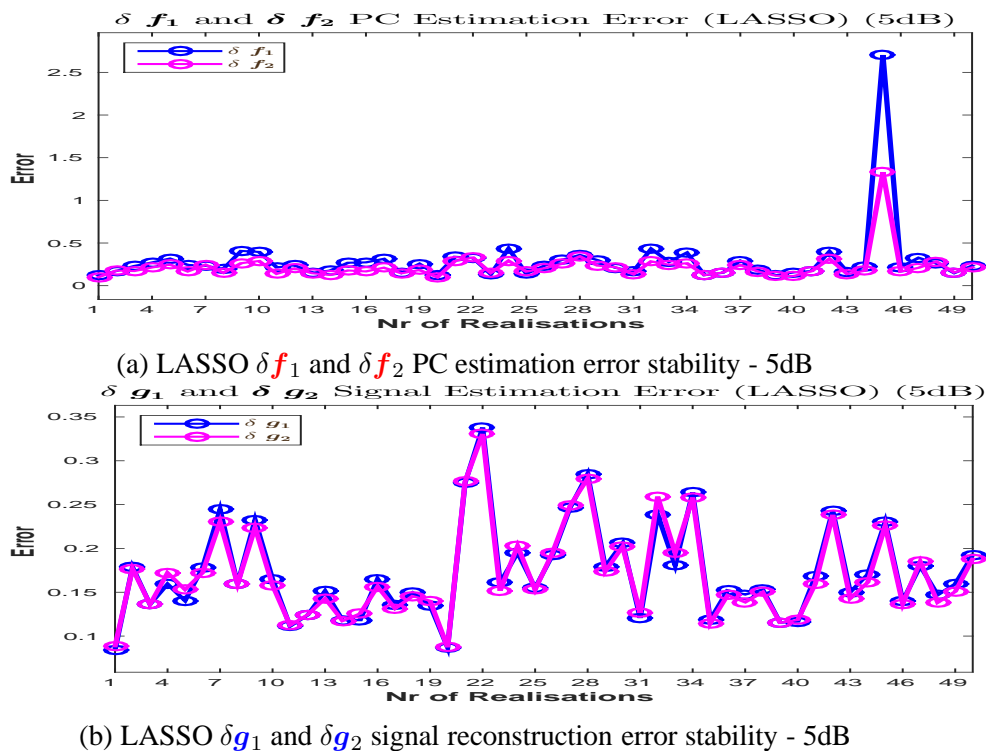
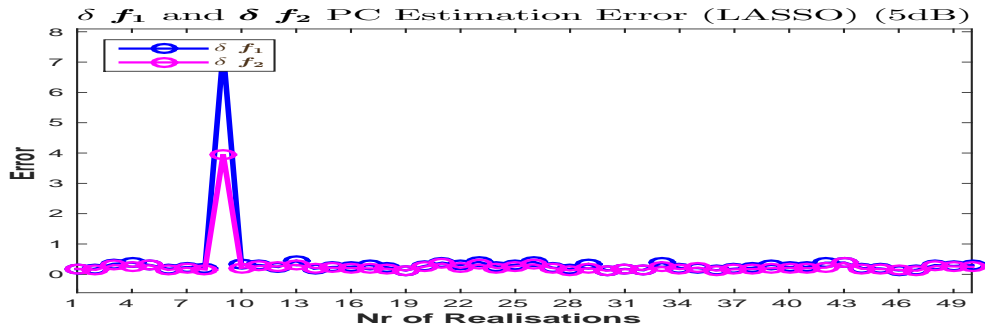


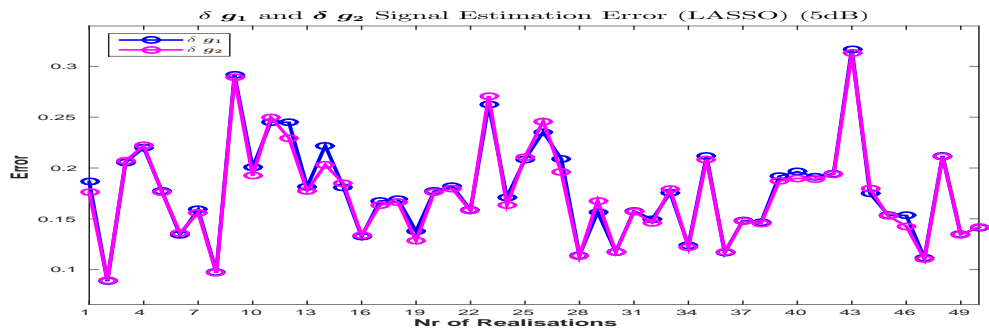
Figure 3.29: LASSO  $L_1$  and  $L_2$  errors stability: PC estimation and signal reconstruction - 5dB

### 3.7 Conclusion and why go to Bayesian

The classical methods presented in this chapter do not give satisfactory results. The FFT approach fails to give the one hour precision for the PC vector, and while via the zero padding

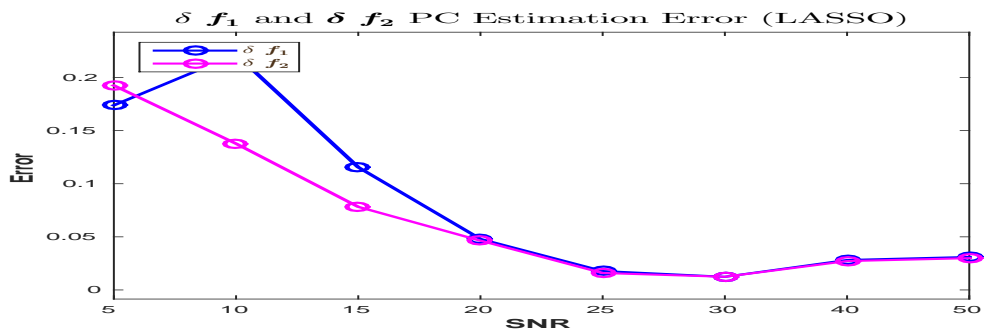


(a) LASSO  $\delta f_1$  and  $\delta f_2$  PC estimation error stability - 5dB

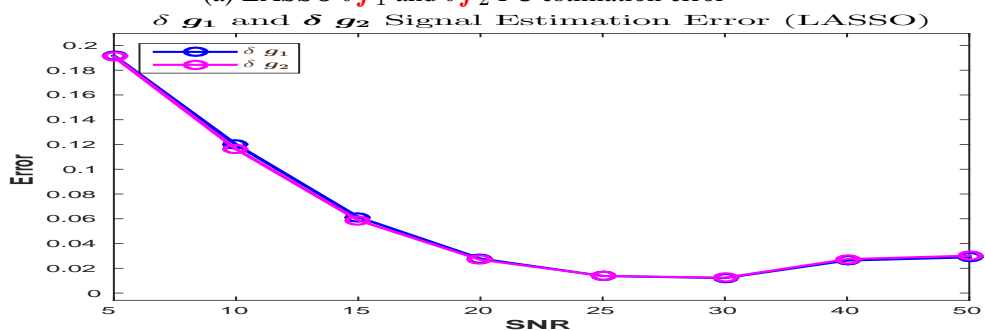


(b) LASSO  $\delta g_1$  and  $\delta g_2$  signal reconstruction error stability - 5dB

Figure 3.30: LASSO  $L_1$  and  $L_2$  errors stability: PC estimation and signal reconstruction - 5dB



(a) LASSO  $\delta f_1$  and  $\delta f_2$  PC estimation error



(b) LASSO  $\delta g_1$  and  $\delta g_2$  signal reconstruction error

Figure 3.31: LASSO  $L_1$  and  $L_2$  errors: PC estimation and signal reconstruction

the precision can be improved, the sparse structure of the PC vector is not achieved. The Direct and Generalized Inversion is extremely sensitive to the level of noise: the estimation of the PC amplitudes vector is very precise in the particular case when the considered signal

has no noise, but fails to correctly estimate the PC amplitudes vector when even very small level of noise is considered, although the reconstruction of the signal is very accurate. Also the Least Squares and the regularized LS fails to correctly estimate the PC amplitudes vector when the input is noisy. Setting the regularization parameter is also critical. Finally while the  $L_1$  regularization can give satisfactory results, even for high levels of noise, this approach has two fundamental drawbacks: the stability of the solution and the determination of the regularization parameter. For the choice of the regularization parameter, there are many methods. We can mention the  $L$ -curve and the cross validation. We used the cross-validation method proposed in the *MATLAB* program for optimal value of the regularization parameter, even if this step has high computational cost. The stability of the solution is a more important drawback: depending on the realization of the noise, the method may not be able to find the right solution, as showed in Figure (3.29) (a). Final point related to all the deterministic regularization methods is that they do not give naturally the possibility of quantifying the uncertainties.

Regardless of their drawbacks in terms of precision or the stability of the solution, the methods presented in this chapter don't allow the estimation of the variances associated with the PC amplitudes vector.

A Bayesian approach allows accounting for the prior information, in our case the sparse structure of the PC amplitude vector, by selecting a sparsity enforcing prior and also allows the estimation of the variances associated with the PC amplitudes vector, by considering the variances in our model as unknowns.

# 4

## Linear Model, Inverse Problem and Bayesian approach with sparsity enforcing prior

### Contents

---

<b>4.1</b>	<b>Forward model and inverse problem approach</b>	<b>52</b>
<b>4.2</b>	<b>General Bayesian inference</b>	<b>55</b>
<b>4.3</b>	<b>Sparsity enforcing priors</b>	<b>56</b>
<b>4.4</b>	<b>Generalized Student-t hierarchical model</b>	<b>58</b>
4.4.1	Generalized Student-t hierarchical prior model	58
4.4.2	Joint MAP Estimation	63
4.4.3	Posterior Mean (via VBA, partial separability)	65
4.4.4	Prior Inverse Gamma parameters setting for $v_{f_j}$	72
4.4.5	Prior Inverse Gamma parameters setting for $v_{\epsilon_i}$	75
4.4.6	Posterior Mean (via VBA, full separability)	77

---

THIS chapter is dedicated to the detailed presentation of the proposed method and the resulting algorithms. First, we present the proposed linear (sinusoidal) model. Other approaches for modelling this problem are mentioned and discussed. Then, the general Bayesian inference is discussed. After that, we give a detailed presentation of the hierarchical model that is build: the probability density functions considered for the likelihood, for the priors and for the hyperparameter priors. The unknowns of the model are estimated from the posterior distribution derived from the hierarchical model, via the Bayes rule, considering Joint Maximum A Posterior and Posterior Mean. For each estimator the computational methods are presented in detail, along with their corresponding algorithm.

## 4.1 Forward model and inverse problem approach

---

The proposed method for improving the precision in PC vector estimation consists, in the first step, in formulating the problem as a linear forward model. This can be done starting from the inverse Fourier transform, which establishes a linear relation between the known signal  $g(t)$  and its PC amplitudes vector. Evidently, in this case, the PC amplitudes vector is the one that is set by default by the FFT method. Considering a matrix that is built in the same way as the one corresponding to the inverse Fourier transform, but using the elements that corresponds to the imposed PC vector, the linear relation between a chronobiological time series and the corresponding PC vector can be described by the following equation:

$$g(t_i) \simeq \sum_{j=1}^M f_1(p_j) \sin(2\pi \frac{t_i}{p_j}) + \sum_{j=1}^M f_2(p_j) \cos(2\pi \frac{t_i}{p_j}), \quad i \in \{1, \dots, N\}, \quad j \in \{1, \dots, M\} \quad (4.1)$$

In Equation (4.1), the following notations are used:

- $t_i, i \in \{1, 2, \dots, N\}$  represents the sampling points of the observed signal; in particular, both for synthetic and real data simulations, we have considered the observed signal sampled every hour. During the real data simulations we had to consider different lengths, depending on the segment that was studied, i.e. 2 days length signals for the segments corresponding to the after treatment segment ( $N = 48$ ), 3 days length signals for the segments corresponding to the before treatment ( $N = 72$ ), 5 days length signals for the segments corresponding to the during treatment ( $N = 120$ ). For the segments that allowed the study of the stability or variability of the PC amplitudes vector, the length of the windows was 4 days. In particular, for the synthetic simulations, the length of the observed signals was considered also 4 days ( $N = 96$ ).
- $g(t_i)$  represents the amplitudes of the observed signal, corresponding to the sampling point  $t_i$
- $p_j, j \in \{1, 2, \dots, M\}$  represents the periods contained by the PC vector; during the simulations, we consider a precision of one hour, in the circadian domain, between 19h and 32h. In order to analyse the behaviour of the potential harmonics corresponding to the periodic components inside the circadian domain, we will also include the periodic components superior to 7h. So, the PC vector considered in the simulations will be  $[8h, 9h, \dots, 32h]$ ,  $M = 25$ .
- $f_1(p_j)$  and  $f_2(p_j)$  represents the amplitudes of PC vectors, corresponding to the period  $p_j$ .

Introducing the notations:

$$g(t_i) = g_i, \quad i \in \{1, \dots, N\} \quad ; \quad f_k(p_j) = f_{kj}, \quad k \in \{1, 2\}, \quad j \in \{1, \dots, M\}, \quad (4.2)$$

equation (4.1) becomes:

$$g_i \simeq \sum_{j=1}^M f_{1j} \sin(2\pi \frac{t_i}{p_j}) + \sum_{j=1}^M f_{2j} \cos(2\pi \frac{t_i}{p_j}) \rightarrow \mathbf{g} \simeq \mathbf{H}_1 \mathbf{f}_1 + \mathbf{H}_2 \mathbf{f}_2 \quad (4.3)$$

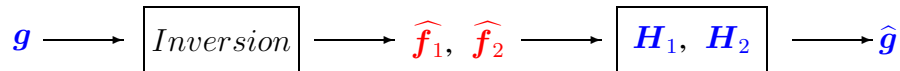
The potential modelling and measurement errors should be accounted. This is critical in our model, since the chronobiological signals are noisy. We want to be able to distinguish between the periodic components corresponding to the biological phenomena and the period components that are due to the noise and uncertainties. The linear model (4.3) is completed by introducing the error vector  $\epsilon$ , leading to the forward linear model:

$$\mathbf{g} = \mathbf{H}_1 \mathbf{f}_1 + \mathbf{H}_2 \mathbf{f}_2 + \epsilon \quad (4.4)$$

where the following notations are used:

- $\mathbf{g}$  represents the observed data i.e. the chronobiological time series:  $\mathbf{g} = [g_1, g_2 \dots g_N]^T \in \mathcal{M}_{N \times 1}$ , a  $N$ -dimensional vector, where the notation  $g(t_i) = g_i$  introduced in (4.2) is used;
- $\mathbf{H}_1$  and  $\mathbf{H}_2$  represent the sine and cosine matrices,  $[\mathbf{H}_1]_{ij} = (\sin(2\pi \frac{t_i}{p_j}))_{ij}$ ,  $[\mathbf{H}_2]_{ij} = (\cos(2\pi \frac{t_i}{p_j}))_{ij} \in \mathcal{M}_{N \times M}$  corresponding to the considered period component vector  $[p_1, p_2, \dots, p_M]^T$ ;
- $\mathbf{f}_1$  and  $\mathbf{f}_2$  represent the unknowns of the forward linear model (4.4), i.e. the periodic components vector amplitudes, corresponding to the sine matrix  $\mathbf{H}_1$  and cosine matrix  $\mathbf{H}_2$ :  $\mathbf{f}_1 = [f_{11}, f_{12}, \dots, f_{1M}]^T$ ,  $\mathbf{f}_2 = [f_{21}, f_{22}, \dots, f_{2M}]^T$ ,  $\in \mathcal{M}_{M \times 1}$ ,  $M$ -dimensional vectors, where the notations  $f_1(p_j) = f_{1j}$  and  $f_2(p_j) = f_{2j}$  introduced in (4.2) are used;
- $\epsilon$  represents the errors:  $\epsilon = [\epsilon_1, \epsilon_2, \dots, \epsilon_N]^T \in \mathcal{M}_{N \times 1}$ , is a  $N$ -dimensional vector; together with the two PC vectors amplitudes,  $\mathbf{f}_1$  and  $\mathbf{f}_2$  the error vector  $\epsilon$  represents the unknowns of the model (4.4).

The main goal is to estimate the unknowns of the linear model (4.4), i.e. the PC amplitude vectors  $\widehat{\mathbf{f}}_1$  and  $\widehat{\mathbf{f}}_2$  and the error vector  $\widehat{\epsilon}$ . In our approach, the strategy used is based on general Bayesian inference: building a hierarchical model, by assigning probability density functions for the likelihood, for the priors, and in order to estimate the hyperparameters that appear in the hierarchical model (namely the variances associated with the unknowns of the linear model (4.4), i.e.  $\widehat{\mathbf{f}}_1$ ,  $\widehat{\mathbf{f}}_2$  and  $\widehat{\epsilon}$ ) for the hyperparameters priors and then estimating the unknowns from the posterior probability density function, obtained via Bayes rule, using the available data  $\mathbf{g}$ :



Once an inversion method is considered and the corresponding estimations are obtained, the estimation accuracy can be verified by comparing the estimates  $\widehat{\mathbf{f}}_1$  and  $\widehat{\mathbf{f}}_2$  with  $\mathbf{f}_1$  and  $\mathbf{f}_2$ . Evidently, such a comparison is possible just in the synthetic case. From the estimates  $\widehat{\mathbf{f}}_1$  and  $\widehat{\mathbf{f}}_2$ , via the linear model (4.4) the reconstructed signal  $\widehat{\mathbf{g}} = \mathbf{H}_1 \widehat{\mathbf{f}}_1 + \mathbf{H}_2 \widehat{\mathbf{f}}_2$  is computed and then compared with the available  $\mathbf{g}$ . This comparison is possible in both the synthetic simulations and the real simulations case. The comparison can be done computing the Normalized Mean Squared Error (NMSE,  $p = 2$ ) or Normalized Mean Absolute Error (NMAE,  $p = 1$ ) via the following relations:

$$\delta_p \mathbf{f} = \frac{\|\mathbf{f} - \widehat{\mathbf{f}}\|_p^p}{\|\mathbf{f}\|_p^p} ; \quad \delta_p \mathbf{g} = \frac{\|\mathbf{g} - \widehat{\mathbf{g}}\|_p^p}{\|\mathbf{g}\|_p^p} , \quad p = 1, 2. \quad (4.5)$$

***Remark on the linear model***

The linear model considered in Equation (4.4) can be expressed in a different form. Introducing the notations

$$\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2] \ ; \ \mathbf{f} = \begin{bmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \end{bmatrix} \quad (4.6)$$

the linear model can be expressed by the following equation:

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\epsilon} \quad (4.7)$$

In particular, Equation (4.7) is expressing an forward model that corresponds to many other application such as signal deconvolution, image restoration, Computed Tomography (CT) image reconstruction, Fourier Synthesis (FS) inversion, microwave imaging [NMD94] and [FDMD07], ultrasound echography, seismic imaging, radio astronomy [KTB04] fluorescence imaging, inverse scattering [CMD97], [FDMD05], [AMD10] and [GHBMD13], Eddy current non destructive testing [NMD96] or SAR imaging [AKZ06]. In all these examples, including the PC amplitude vector estimation, the common inverse problem is to estimate  $\mathbf{f}$  from the observations of  $\mathbf{g}$ . In general, the inverse problems are ill-posed [Had01], since the conditioning number of the matrix  $\mathbf{H}$  is very high. This means that, in practice, the data  $\mathbf{g}$  alone is not sufficient to define an unique and satisfactory solution.

One particular case of the linear model, Equation (4.7) is when the error vector is neglected ( $\boldsymbol{\epsilon} = \mathbf{0}$ ). For this particular case, the linear model is given by the equation

$$\mathbf{g} = \mathbf{H}\mathbf{f}. \quad (4.8)$$

If the matrix  $\mathbf{H}$  was invertible and orthogonal, i.e.  $\mathbf{H}^T\mathbf{H} = \mathbf{I}$  (this is the case of the Fast Fourier Transform for the particular case when  $M = N$ ), the solution is given by

$$\hat{\mathbf{f}} = \mathbf{H}^T\mathbf{g}. \quad (4.9)$$

If the matrix is just invertible but not orthogonal we could write

$$\hat{\mathbf{f}} = \mathbf{H}^{-1}\mathbf{g}. \quad (4.10)$$

If the matrix is not invertible we can still define generalized inverse or pseudo-inverse solutions. The Minimum Norm Solution (MNS) and the Least Square Solution (LSS) are two examples:

$$\text{MNS} : \hat{\mathbf{f}}_{MN} = (\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{H}^T\mathbf{g}, \quad M < N \quad (4.11)$$

$$\text{LSS} : \hat{\mathbf{f}}_{LS} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{g}, \quad M > N \quad (4.12)$$

However, neither solutions Equation (4.10), Equation (4.11) and Equation (4.12) are helpful in practice, since they don't account for the noise. These inverse solutions are in general too sensitive to the errors due to the ill-conditioning of the matrix  $\mathbf{H}$  in Equation (4.10),  $\mathbf{H}^T\mathbf{H}$  in Equation (4.11) and  $\mathbf{H}\mathbf{H}^T$  in Equation (4.12).

Among the techniques that have been successfully used for solving Equation (4.7) are regularization theory and the Bayesian inversion. Quadratic and Tikhonov regularization

is developed in [DR85], [Hol92], [PL99] and [YLZF12]. Total Variation (TV) is developed in [CW98], [ZL11] and [YFZ12], different entropy based regularization are developed in [Wig77], [Dee81], [Don81] and [MDD88],  $L_p$  and sparsity enforcing in [YBW85] and [WHmPw08], blind deconvolution and applications in [NPQ00], [JSM<sup>+</sup>02], [Fio04] and [MSSM11]. Cross Validation (CV) and generalized CV methods for determining the regularization parameter are developed in [Wah82] and [LN11], Bernouilli-Gaussian models in [GDI90] and [CGI96], the Compress Sensing (CS) approach in [RMW12], document image restoration in [PB13] and [LB13], joint restoration and segmentation in [AMD09].

Considering the model (4.4) or equivalent (4.7) we focus on an inversion based on General Bayesian Inference, building an hierarchical model, based on the generalized Student-t prior distribution modelling the PC amplitude vectors in order to enforce the sparsity. Accounting for sparsity has been considered in many ways. One way is via regularization techniques, by using  $L_0$  or  $L_1$  norms. This approach was developed in [DDM04], [Tro06], [TGS06], [CWB08] and [PS10]. One of these methods which has become now the standard is LASSO [Tip01]. Therefore, during the simulations chapter, we will consider LASSO for comparisons with the proposed algorithm. One other way is via Bayesian inference using strict sparsity or sparsity enforcing priors. For the strict sparsity requirement, very often Bernouilli distribution is used. For example, Bernouilli-Gaussian [GDI90], [CGI96] and [GILC08], Bernouilli-Laplace [DB08] or Bernouilli-Gamma, while for the sparsity enforcing, mainly three categories of priors have been considered and used very often: Generalized Gaussian (GG), Mixture models and heavy tailed probability laws such as Student-t. A review of these priors is discussed in [Tip01] and [MD12] and their references.

Furthermore, the hierarchical model developed in this chapter will account for the non stationarity of the noise. For this, in literature zero mean Gaussian model with unknown varying variance has been considered in [CGHM14] and a Cauchy-Gaussian model in [CKS15]. Like the prior distribution modelling the PC amplitude vectors, the Student-t distribution is used for modelling the non-stationarity.

## 4.2 General Bayesian inference

In Section (4.1) we have described the strategy adopted for doing the inversion in Equation (4.4) or equivalent (4.7): building an hierarchical model enforcing the sparsity for the PC amplitudes vectors and taking into account the non stationarity of the noise, based on general Bayesian inference. Considering the linear inverse problem of Equation (4.7), the Bayesian inference starts by the fundamental relation given by the Bayes rule:

$$p(\mathbf{f}|\mathbf{g}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \frac{p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f}|\boldsymbol{\theta}_2)}{p(\mathbf{g}|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2)}, \boldsymbol{\theta} = (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \quad (4.13)$$

where  $\boldsymbol{\theta}$  represents the hyperparameters that appear in the model, namely the variances associated with the unknowns of our model, i.e the PC amplitudes vector and the noise vector  $\boldsymbol{\epsilon}$ .

The Bayes rule can be interpreted as a proportionality relation between the posterior law and the product of the prior law (the sparsity information) and the likelihood:

$$p(\mathbf{f}|\mathbf{g}, \boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \propto p(\mathbf{g}|\mathbf{f}, \boldsymbol{\theta}_1) p(\mathbf{f}|\boldsymbol{\theta}_2) \quad (4.14)$$



## CHAPTER 4. LINEAR MODEL, INVERSE PROBLEM AND BAYESIAN APPROACH WITH SPARSITY ENFORCING PRIOR

---

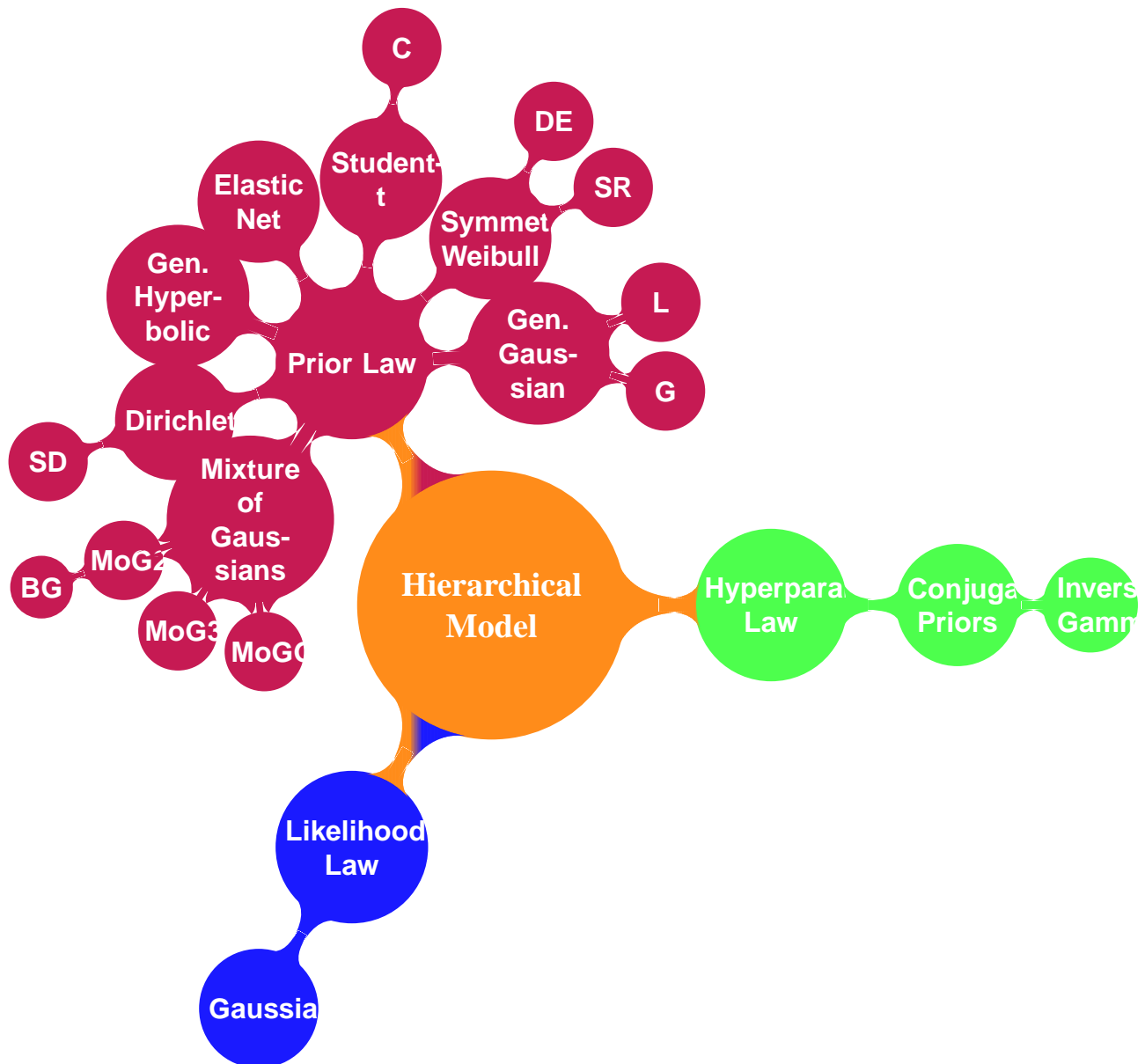
The probabilistic model that assigns particular distributions for the prior and likelihood laws represents the hierarchical model. An extension of Equation (4.14) is the general Bayesian Inference, where the hyperparameters involved in the Bayes Rule,  $\theta = (\theta_1, \theta_2)$  are considered unknowns and are also estimated from the posterior law:

$$p(\mathbf{f}, \theta_1, \theta_2 | \mathbf{g}) \propto p(\mathbf{g} | \mathbf{f}, \theta_1) p(\mathbf{f} | \theta_2) p(\theta_1) p(\theta_2) \quad (4.15)$$

### 4.3 Sparsity enforcing priors

---

In the Bayesian framework, distributions must be assigned for the likelihood (derived from the prior distribution assigned for modelling the errors of the forward model), for the prior (in our case one that allows enforcing the sparsity) and also for the hyperparameters of the model (i.e. the variances associated with the PC amplitudes vector and with the errors). In literature, [Tip01] and [MD12] and their references certain classes of distribution (heavy-tailed, mixture models) are well known as good sparsity enforcing priors. The overview is presented in the following scheme:



Prior models that can be considered are: **a)** Generalized Gaussian (GG) with Gaussian (G) and Laplace; **b)** Weibull (W) with Symmetric Rayleigh (SR) or Double Exponential (DE) as particular cases; **c)** Student-t (St) with Cauchy (C) as particular case; **d)** Elastic net; **e)** Generalized Hyperbolic model; **f)** Dirichlet and with Symmetric Dirichlet (SD) as particular case; **g)** Mixture of Gaussians, with **(i)** Mixture of two centred Gaussians (MoG2), one with very small and one with a large variance, with Bernoulli-Gauss (BG) (also called Spike and slab) as a particular case, **(ii)** Mixture of three Gaussians (MoG3), one centred with very small variance and two symmetrically centred on positive and negative axes and large variances; **(iii)** Mixture of one Gaussian and one or two Gammas (MoG-Gammas), and **h)** Bernoulli-Multinomial (BMult) or mixture of Dirichlet (MoD).

Mainly, the prior models are classified in three categories: Generalized Gaussian, heavy tailed distributions and Mixture Models, which result to hierarchical models.

## 4.4 Generalized Student-t hierarchical model

---

This section presents the assignment of the probability laws corresponding to the likelihood, the prior and the prior corresponding to the hyperparameters, i.e. the variances associated with the PC amplitude vectors and the error vector. The goal is to build a hierarchical model that accounts for the prior informations, i.e. the sparsity of the PC amplitude vectors and the non-stationary noise. To account for both of these two prior informations, we model the noise  $\epsilon$  as a zero mean non stationary Gaussian with unknown variances on which we assign Inverse Gamma priors, and to enforce the sparsity, we use the generalized Student-t distribution as the prior modelling the PC amplitude vectors  $f_1$  and  $f_2$ , expressed as the marginal of the bivariate Normal - Inverse Gamma distribution. The main advantage of the Student-t distribution is that, thanks to its Infinite Gaussian Scaled Mixture (IGSM) property, it can be used in a hierarchical Gaussian - Gamma or Gaussian - Inverse Gamma model. In this way, in fact both for the non stationarity of the noise and for sparsity enforcing we have the same prior model structure: Gaussian with unknown variances on which we assign Inverse Gamma priors. The Bayesian framework with different priors both on the noise and on the solution is a very well known approach in the literature, and it goes back to 1950, using the Normal distribution [HW83] or the Poisson distribution [MSC99] for the noise and the Normal distribution for the solution. More specific priors and in particular the Markovian model [GG84] and [IG90], non Gaussian priors [NMD94], the hierarchical models [GLRI93] are more recent. The main difficulties in those methods have been more on the computational aspects. Beside the classical Gaussian approximation [RMC09] and the MCMC methods [GRS96],[DDJ09],[GILC08], we may mention the more recent ones: the Approximate Bayesian Computation (ABC) [BCMR09],[Blu10],[BF10],[MPRR11], Variational Bayesian Approximation (VBA) [KMS<sup>+</sup>12b],[KMS<sup>+</sup>12a] and Message Passing (MP) [Bea03],[WBJ05],[PCS11],[SR12] methods.

In general, a hierarchical model represents the set of probability laws assigned for the probabilities involved in (4.15), namely the assignment of the likelihood  $p(g|f, \theta_1)$ , the prior  $p(f|\theta_2)$  and the hyperparameters priors  $p(\theta_1)$  and  $p(\theta_2)$ . The equations corresponding to each probability density function assigned in the model are presented and discussed in detail in Subsection (4.4.1). The unknowns of the hierarchical prior model are estimated from the expression of the posterior distribution, via the Bayes rule, Equation (4.15) using two different estimation techniques: Joint Maximum A Posteriori (JMAP) and PM (Posterior Mean), via Variational Bayesian Approximation (VBA). The computations corresponding to those estimations and the resulting algorithm are detailed in Subsection (4.4.2) and Subsection (4.4.3).

### 4.4.1 Generalized Student-t hierarchical prior model

In the first step, the error vector  $\epsilon$  of the linear model (4.4) is modelled. We propose to use a non-stationary Gaussian model for the conditional probability  $p(\epsilon_i|v_{\epsilon_i})$ , i.e. every point in the error vector conditioned by the corresponding variance,  $\epsilon_i|v_{\epsilon_i}$  is considered as a zero mean normal distribution with the corresponding variance:

$$p(\epsilon_i|v_{\epsilon_i}) = \mathcal{N}(\epsilon_i|0, v_{\epsilon_i}), \quad i \in \{1, 2, \dots, N\}, \quad (4.16)$$

where the corresponding variances  $v_{\epsilon_i}$  are considered to be unknowns. For having the possibility to estimate them we model them as Inverse Gamma distributions:

$$p(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}), \quad i \in \{1, 2, \dots, N\}, \quad (4.17)$$

where  $\alpha_{\epsilon_0}$  and  $\beta_{\epsilon_0}$  represent the shape and scale parameters corresponding to the Inverse Gamma distribution associated with the corresponding variance of every point in the error vector,  $v_{\epsilon_i}, i \in \{1, 2, \dots, N\}$ . Doing this, the joint probability of the error vector points and their corresponding variances  $(\epsilon_i, v_{\epsilon_i}), i \in \{1, 2, \dots, N\}$  are modelled as a bivariate Normal - Inverse Gamma distribution:

$$p(\epsilon_i, v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \mathcal{N}(\epsilon_i|0, v_{\epsilon_i})\mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}), \quad i \in \{1, 2, \dots, N\}, \quad (4.18)$$

and the probability of each error vector point  $\epsilon_i$  is modelled by its marginal, i.e. a generalized Student-t distribution:

$$p(\epsilon_i|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \int \mathcal{N}(\epsilon_i|0, v_{\epsilon_i})\mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) \, dv_{\epsilon_i}, \quad i \in \{1, 2, \dots, N\}, \quad (4.19)$$

More general, the equations introduced in (4.16) and (4.17), can be used to write:

$$\begin{cases} p(\boldsymbol{\epsilon}|\mathbf{v}_\epsilon) = \mathcal{N}(\boldsymbol{\epsilon}|\mathbf{0}, \mathbf{V}_\epsilon) \\ p(\mathbf{v}_\epsilon|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \prod_{i=1}^N \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) \end{cases} \quad (4.20)$$

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

$$\mathbf{v}_\epsilon = [v_{\epsilon_1} \dots v_{\epsilon_i} \dots v_{\epsilon_N}]^T \quad ; \quad \mathbf{V}_\epsilon = \text{diag}[\mathbf{v}_\epsilon] \quad (4.21)$$

The likelihood  $p(\mathbf{g}|\mathbf{f}, \mathbf{v}_\epsilon)$  is obtained considering the linear model, Equation (4.4) and the assigned distribution for the error vector  $\boldsymbol{\epsilon}$  conditioned by the variance  $\mathbf{v}_\epsilon$ ,  $p(\boldsymbol{\epsilon}|\mathbf{v}_\epsilon)$ , Equation (4.20). The distribution modelling the likelihood is also a multivariate Normal distribution, having the same covariance matrix  $\mathbf{V}_\epsilon$  and the mean  $\mathbf{H}_1\mathbf{f}_1 + \mathbf{H}_2\mathbf{f}_2$ , derived from the linear model Equation (4.4):

$$p(\mathbf{g}|\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon) = \mathcal{N}(\mathbf{g}|\mathbf{H}_1\mathbf{f}_1 + \mathbf{H}_2\mathbf{f}_2, \mathbf{V}_\epsilon) \quad (4.22)$$

The likelihood and the prior distribution corresponding to the noise variance  $\mathbf{v}_\epsilon$  represents the first part of the hierarchical model:

$$\begin{cases} p(\mathbf{g}|\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon) = \mathcal{N}(\mathbf{g}|\mathbf{H}_1\mathbf{f}_1 + \mathbf{H}_2\mathbf{f}_2, \mathbf{V}_\epsilon) \\ p(\mathbf{v}_\epsilon|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \prod_{i=1}^N \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) \end{cases} \quad (4.23)$$

The proposed prior distribution is a generalized Student-t distribution  $St_g$ , in order to enforce the sparsity, obtained via a Normal - Inverse Gamma distribution. A generalization of the Student-t distribution is obtained from the Normal - Inverse Gamma distribution by not introducing any supplementary conditions for the shape and scale parameters corresponding to the Inverse Gamma. This generalization is discussed in detail in Subsection (4.4.4). The probability density function of the standard Student-t distribution is given in Equation (4.24):

$$St(x|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad (4.24)$$

#### 4.4.1 - Generalized Student-t hierarchical prior model

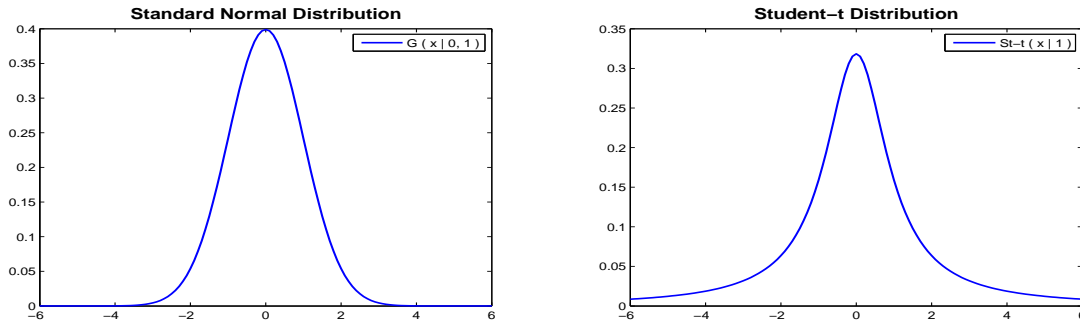


Figure 4.1: Normal and Student-t distribution

The Student-t distribution with one degree of freedom ( $\nu = 1$ ) (Cauchy distribution) and the standard Normal distribution are presented in Figure (4.1). The Student-t distribution is a good sparsity enforcing prior due to the heavy tailed structure: the tail of the distribution is above the tail of the Normal distribution, allowing certain cases of high values i.e. the non zero values in the periodic component vector amplitudes, corresponding to the biological periods expressed in the chronobiological signal. A comparison between the standard Normal distribution and the Student-t distribution ( $\nu = 1$ ) is presented in Figure (4.2). A

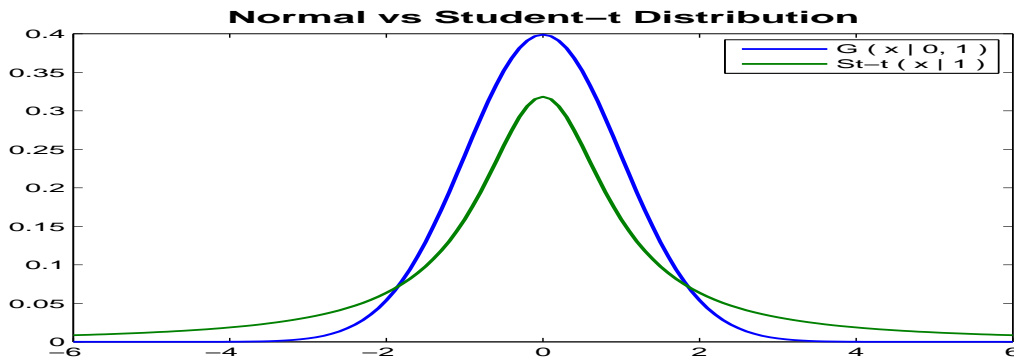


Figure 4.2: Normal vs. Student-t distribution

direct assignment of Student-t distributions for the prior laws  $p(\mathbf{f}_1)$  and  $p(\mathbf{f}_2)$  leads to a non-quadratic criterion when estimating  $\mathbf{f}_1$ ,  $\mathbf{f}_2$ . The Student-t distributions corresponding to the prior distributions modelling the PC amplitudes vectors  $\mathbf{f}_1$  and  $\mathbf{f}_2$  can be assigned via Normal - Inverse Gamma, jointly modelling the periodic component vectors amplitudes and their corresponding variances. For the variances of  $\mathbf{f}_1$  and  $\mathbf{f}_2$  we assume general models. Also, the considered model will assign the same variances for  $\mathbf{f}_1$  and  $\mathbf{f}_2$ , linking the two periodic component vectors:

$$\mathbf{v}_f = [v_{f_1} \dots v_{f_j} \dots v_{f_M}]^T ; \mathbf{V}_f = \text{diag}[\mathbf{v}_f] \quad (4.25)$$

Therefore, for sparsity enforcing for the PC amplitudes vectors, the joint distributions corresponding to the PC amplitudes vectors elements and their corresponding variances are modelled via the bivariate Normal - Inverse Gamma distribution, imposing equal values for the shape and scale parameter of the Inverse Gamma distribution,  $\alpha_{f0} = \beta_{f0} = \frac{\nu_f}{2}$ :

$$(\mathbf{f}_{kj}, v_{f_j}) \sim \mathcal{N}(\mathbf{f}_{kj} | 0, v_{f_j}) \mathcal{IG}(v_{f_j} | \frac{\nu_f}{2}, \frac{\nu_f}{2}), k \in \{1, 2\}, j \in \{1, 2, \dots, M\} \quad (4.26)$$

Doing so, the marginal of the joint distribution, modelling the PC amplitudes vectors elements is a Student-t distribution, Equation (4.24):

$$p(\mathbf{f}_{kj}|\nu_f) = \int \mathcal{N}(\mathbf{f}_{kj}|0, \mathbf{v}_{f_j}) \mathcal{IG}(\mathbf{v}_{f_j}|\frac{\nu_f}{2}, \frac{\nu_f}{2}) d\mathbf{v}_{f_j} = \mathcal{St}(\mathbf{f}_{kj}|\nu_f) \quad (4.27)$$

This approach can be extended by considering a Normal - Inverse Gamma distribution for the joint distribution  $(\mathbf{f}_{kj}, \mathbf{v}_{f_j})$ , without imposing the equality  $\alpha_{f_0} = \beta_{f_0} = \frac{\nu_f}{2}$  in the expression of the Inverse Gamma distribution. Considering  $\alpha_{f_j,0}, \beta_{f_j,0}$  as the parameters of the Inverse Gamma distribution in the Normal - Inverse Gamma, the joint probability of  $(\mathbf{f}_{kj}, \mathbf{v}_{f_j})$  is modelled as:

$$(\mathbf{f}_{kj}, \mathbf{v}_{f_j}) \sim \mathcal{N}(\mathbf{f}_{kj}|0, \mathbf{v}_{f_j}) \mathcal{IG}(\mathbf{v}_{f_j}|\alpha_{f_0}, \beta_{f_0}), k \in \{1, 2\}, j \in \{1, 2, \dots, M\} \quad (4.28)$$

The marginal modelling the elements of the PC amplitudes vectors is:

$$p(\mathbf{f}_{kj}|\alpha_{f_0}, \beta_{f_0}) = \frac{\Gamma(\alpha_{f_0} + \frac{1}{2})}{\sqrt{2\beta_{f_0}\pi} \Gamma(\alpha_{f_0})} \left(1 + \frac{\mathbf{f}_{kj}^2}{2\beta_{f_0}}\right)^{-(\alpha_{f_0} + \frac{1}{2})} = \mathcal{St}_g(\mathbf{f}_{kj}|\alpha_{f_j,0}, \beta_{f_j,0}) \quad (4.29)$$

The Normal - Inverse Gamma distribution, modelling  $(\mathbf{f}_{kj}, \mathbf{v}_{f_j})$  and the marginal modelling  $\mathbf{f}_{kj}$  are presented in Figure (A.1). The two parameters  $\alpha_{f_0}, \beta_{f_0}$  introduced for defining the

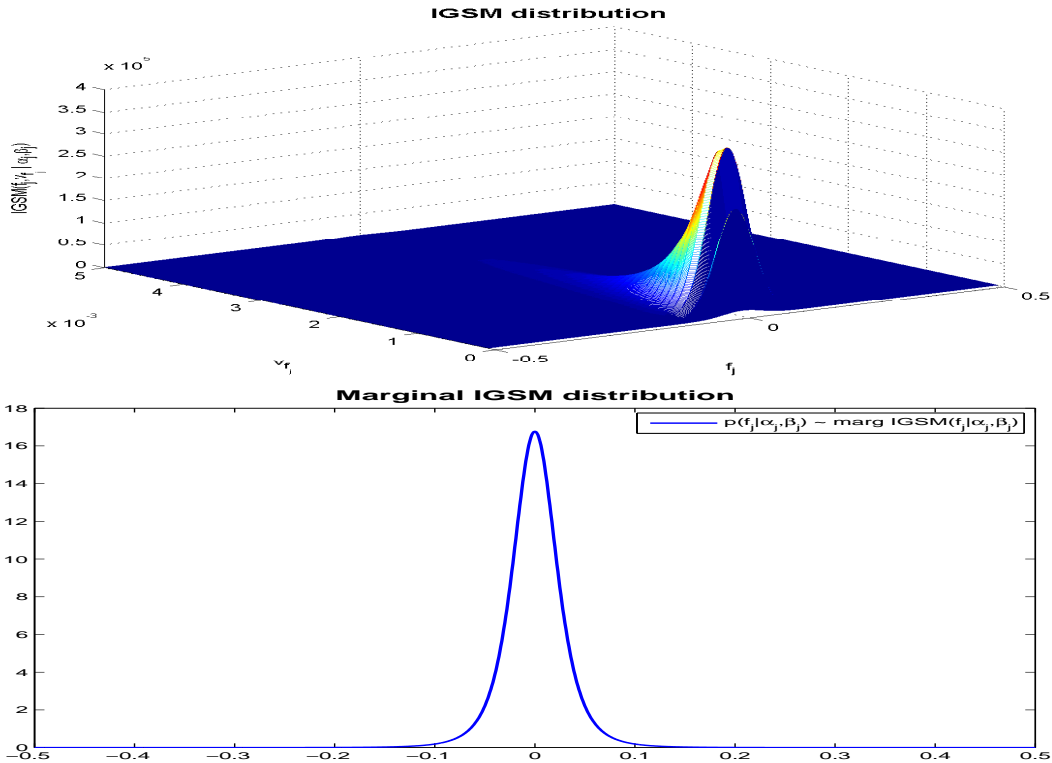


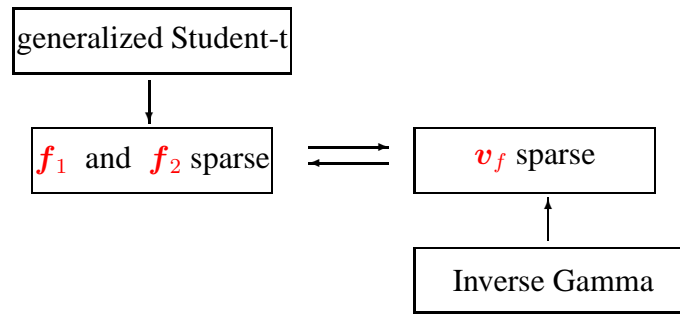
Figure 4.3: Normal - Inverse Gamma distribution  $(\mathbf{f}_{kj}, \mathbf{v}_{f_j})$  and the corresponding marginal  $\mathbf{f}_{kj}|\alpha_{f_0}, \beta_{f_0}$

Normal - Inverse Gamma joint distribution are directly modelling the variances  $\mathbf{v}_{f_j}$  via the Inverse Gamma prior distribution  $\mathcal{IG}(\mathbf{v}_{f_j}|\alpha_{f_0}, \beta_{f_0})$  and also modelling  $\mathbf{f}_{kj}$  via equation (A.25). In order to enforce the sparsity, the two parameters must be chosen such that the Normal

#### 4.4.1 - Generalized Student-t hierarchical prior model

- Inverse Gamma marginal is a heavy tailed distribution and such that the behaviour of the variance is in accordance with the sparse structure of the  $f_{kj}$ .

In fact, in such an approach, the mechanism of sparsity is based not only on the heavy tailed property of the prior distribution, but also on a particular behaviour of the variances. In such an approach a bivariate prior is set for the unknown of the model that needs to be estimated (the PC amplitudes vectors in our case) and for the corresponding variance. The algorithm that results is an iterative one, updating at every iteration both the unknowns of the model (PC amplitudes vectors  $f_1$  and  $f_2$ ) and the corresponding variance. In order to obtain a sparse solution for the unknowns, the structure of the variance must be sparse itself. In particular the variances associated with the zero or close to zero points from the unknown of the model must be small, and the variances associated with the non-zero elements of the sparse unknowns of the model must be significant.



Therefore, the shape and scale parameters of the Inverse Gamma distribution must be chosen such that the variance vector  $v_f$  is sparse, i.e. the expected value of the elements  $v_{f_j}$  is close to zero,  $E_{IG} [v_{f_j}] \searrow 0$ . Furthermore, in a Bayesian approach, we may have a prior knowledge for the numerical value associated with the variance of the Inverse Gamma modelling  $v_{f_j}$ , i.e.  $\text{Var}_{IG} [v_{f_j}] = w$ , where  $w$  is the numerical value obtained via prior knowledge. Evidently, depending on the shape and scale parameters corresponding to the Inverse Gamma distribution, the behaviour of the marginal corresponding to the Normal - Inverse Gamma bivariate distribution, i.e. the generalized Student-t distribution, Equation (A.25) is different. Figure (4.4) presents a comparison between the standard Normal distribution and the generalized Student-t distribution for different values of the two parameters. We note that in order to select a prior model that enforces sparsity, the two parameters  $\alpha_{f_j0}$  and  $\beta_{f_j0}$  must be chosen such that the generalized Student-t distribution modelling the PC amplitudes vectors is concentrated around zero, i.e. it's variance is very small,  $\text{Var}_{St_g} [f_{kj}] \searrow 0$ .

For modelling the priors corresponding to  $f_{kj}$ ,  $k \in \{1, 2\}$ ,  $j \in \{1, 2, \dots, M\}$  as in (A.25), we assign Normal - Inverse Gamma distributions for the joint priors  $(f_{1j}, v_{f_j})$  and  $(f_{2j}, v_{f_j})$ . Therefore,  $f_{1j}|v_{f_j}$  and  $f_{2j}|v_{f_j}$  are modelled as zero mean Normal distributions, and the variances  $v_{f_j}$  are modelled as Inverse Gamma distributions, for  $j \in \{1, 2, \dots, M\}$ . In this way, the vector  $f_1$  conditioned by the variance vector  $v_f$  is modelled by a zero mean multivariate Normal distribution, with the covariance matrix defined in (4.25), the variance vector  $v_f$  is modelled as a product of Inverse Gamma distributions and the vector  $f_2$  conditioned by the variance vector  $v_f$  is modelled by a zero mean multivariate Normal distribution, with the covariance matrix defined in (4.25):

$$\begin{cases} p(\mathbf{f}_1|\mathbf{v}_f) = \mathcal{N}(\mathbf{f}_1|\mathbf{0}, \mathbf{V}_f) \\ p(\mathbf{f}_2|\mathbf{v}_f) = \mathcal{N}(\mathbf{f}_2|\mathbf{0}, \mathbf{V}_f) \\ p(\mathbf{v}_f|\alpha_{f0}, \beta_{f0}) = \prod_{j=1}^M \mathcal{IG}(v_{f_j}|\alpha_{f0}, \beta_{f0}) \end{cases} \quad (4.30)$$

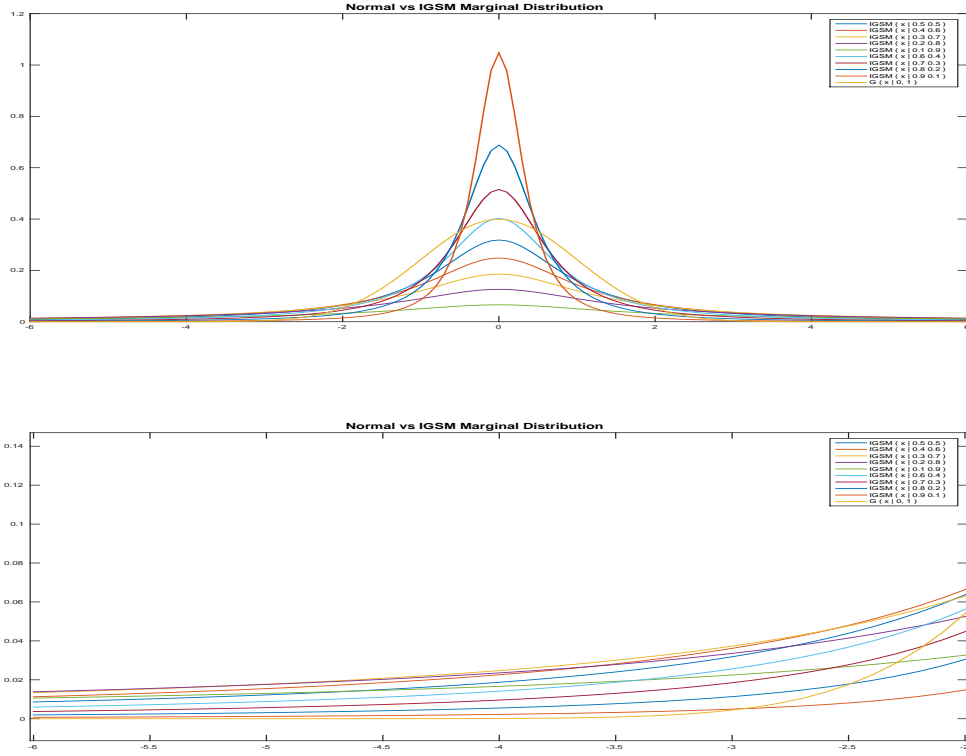


Figure 4.4: Normal vs. Normal - Inverse Gamma marginal distributions

The likelihood and the error variance prior (4.23) and the prior (4.30) represents the IGSM Hierarchical Model. The analytical form of the Hierarchical Model written via proportionalities is the following:

$$\left\{ \begin{array}{l} p(\mathbf{g}|\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon) \propto |\mathbf{V}_\epsilon|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \|\mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 \right\} \\ p(\mathbf{v}_\epsilon|\alpha_{\epsilon 0}, \beta_{\epsilon 0}) \propto \prod_{i=1}^N \frac{\beta_{\epsilon 0}^{\alpha_{\epsilon 0}}}{\Gamma(\alpha_{\epsilon 0})} v_{\epsilon_i}^{-(\alpha_{\epsilon 0}+1)} \exp \left\{ -\beta_{\epsilon 0} v_{\epsilon_i}^{-1} \right\} \\ p(\mathbf{f}_1|\mathbf{v}_f) \propto |\mathbf{V}_f|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1\|^2 \right\} \\ p(\mathbf{f}_2|\mathbf{v}_f) \propto |\mathbf{V}_f|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_2\|^2 \right\} \\ p(\mathbf{v}_f|\alpha_{f 0}, \beta_{f 0}) \propto \prod_{j=1}^M \frac{\beta_{f 0}^{\alpha_{f 0}}}{\Gamma(\alpha_{f 0})} v_{f_j}^{-(\alpha_{f 0}+1)} \exp \left\{ -\beta_{f 0} v_{f_j}^{-1} \right\} \end{array} \right. \quad (4.31)$$

From the Hierarchical Model the posterior distribution can be obtained via the proportionality relation considered in (4.15):

$$p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f|\mathbf{g}) \propto p(\mathbf{g}|\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon) p(\mathbf{v}_\epsilon|\alpha_{\epsilon 0}, \beta_{\epsilon 0}) p(\mathbf{f}_1|\mathbf{v}_f) p(\mathbf{f}_2|\mathbf{v}_f) p(\mathbf{v}_f|\alpha_{f 0}, \beta_{f 0}) \quad (4.32)$$

#### 4.4.2 Joint MAP Estimation

This section presents the results obtained by the estimation of the unknowns from the posterior distribution, Equation (4.32), via Joint Maximum A Posteriori (JMAP), a point estimator of



#### 4.4.2 - Joint MAP Estimation

the unobserved quantities  $\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f$  on the basis of the available data  $\mathbf{g}$ . JMAP estimator is defined as:

$$\left(\widehat{\mathbf{f}}_1, \widehat{\mathbf{f}}_2, \widehat{\mathbf{v}}_\epsilon, \widehat{\mathbf{v}}_f\right) = \underset{(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\operatorname{arg\,max}} p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) = \underset{(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\operatorname{arg\,min}} \mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f), \quad (4.33)$$

where for the second equality the criterion  $\mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)$  is defined as:

$$\mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f) = -\ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \quad (4.34)$$

The MAP estimator is the solution minimizing the criterion  $\mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)$ . One of the simplest optimisation algorithm that we can use is an alternate optimization with respect to the each unknown:

- With respect to  $\mathbf{f}_1$ :

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\partial \mathbf{f}_1} = 0 &\Leftrightarrow \frac{\partial}{\partial \mathbf{f}_1} \left( \|\mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 + \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1\|^2 \right) = 0 \\ &\Leftrightarrow -\mathbf{H}_1^T \mathbf{V}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2) + \mathbf{V}_f^{-1} \mathbf{f}_1 = 0 \\ &\Leftrightarrow [\mathbf{H}_1^T \mathbf{V}_\epsilon^{-1} \mathbf{H}_1 + \mathbf{V}_f^{-1}] \mathbf{f}_1 = \mathbf{H}_1^T \mathbf{V}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}_2 \mathbf{f}_2) \\ &\Rightarrow \widehat{\mathbf{f}}_1 = [\mathbf{H}_1^T \mathbf{V}_\epsilon^{-1} \mathbf{H}_1 + \mathbf{V}_f^{-1}]^{-1} \mathbf{H}_1^T \mathbf{V}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}_2 \mathbf{f}_2) \end{aligned}$$

- With respect to  $\mathbf{f}_2$ :

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\partial \mathbf{f}_2} = 0 &\Leftrightarrow \frac{\partial}{\partial \mathbf{f}_2} \left( \|\mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 + \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_2\|^2 \right) = 0 \\ &\Leftrightarrow -\mathbf{H}_2^T \mathbf{V}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2) + \mathbf{V}_f^{-1} \mathbf{f}_2 = 0 \\ &\Leftrightarrow [\mathbf{H}_2^T \mathbf{V}_\epsilon^{-1} \mathbf{H}_2 + \mathbf{V}_f^{-1}] \mathbf{f}_2 = \mathbf{H}_2^T \mathbf{V}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1) \\ &\Rightarrow \widehat{\mathbf{f}}_2 = [\mathbf{H}_2^T \mathbf{V}_\epsilon^{-1} \mathbf{H}_2 + \mathbf{V}_f^{-1}]^{-1} \mathbf{H}_2^T \mathbf{V}_\epsilon^{-1} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1) \end{aligned}$$

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

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\partial v_{\epsilon_i}} = 0 &\Leftrightarrow \frac{\partial}{\partial v_{\epsilon_i}} \left[ \left( \alpha_{\epsilon_i 0} + \frac{3}{2} \right) \ln v_{\epsilon_i} + \right. \\ &\quad \left. + \left( \beta_{\epsilon_i 0} + \frac{1}{2} (g_i - \mathbf{H}_{2i} \mathbf{f}_1 - \mathbf{H}_{2i} \mathbf{f}_2)^2 \right) v_{\epsilon_i}^{-1} \right] = 0 \\ &\Leftrightarrow \left( \alpha_{\epsilon_i 0} + \frac{3}{2} \right) v_{\epsilon_i} - \left( \beta_{\epsilon_i 0} + \frac{1}{2} (g_i - \mathbf{H}_{2i} \mathbf{f}_1 - \mathbf{H}_{2i} \mathbf{f}_2)^2 \right) = 0 \\ &\Rightarrow \widehat{v}_{\epsilon_i} = \frac{\beta_{\epsilon_i 0} + \frac{1}{2} (g_i - \mathbf{H}_{2i} \mathbf{f}_1 - \mathbf{H}_{2i} \mathbf{f}_2)^2}{\alpha_{\epsilon_i 0} + \frac{3}{2}} \end{aligned}$$

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

$$\begin{aligned} \frac{\partial \mathcal{L}(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f)}{\partial v_{f_j}} = 0 &\Leftrightarrow \frac{\partial}{\partial v_{f_j}} \left( [\alpha_{f_0} + 2] \ln v_{f_j} + \left[ \beta_{f_0} + \frac{1}{2} (f_{1j}^2 + f_{2j}^2) \right] v_{f_j}^{-1} \right) = 0 \\ &\Leftrightarrow (\alpha_{f_0} + 2) v_{f_j} - \left( \beta_{f_0} + \frac{1}{2} (f_{1j}^2 + f_{2j}^2) \right) = 0 \\ &\Rightarrow \widehat{v}_{f_j} = \frac{\beta_{f_0} + \frac{1}{2} (f_{1j}^2 + f_{2j}^2)}{\alpha_{f_0} + 2} \end{aligned}$$

The scheme of the iterative algorithm obtained via JMAP estimation is presented Figure (4.5):

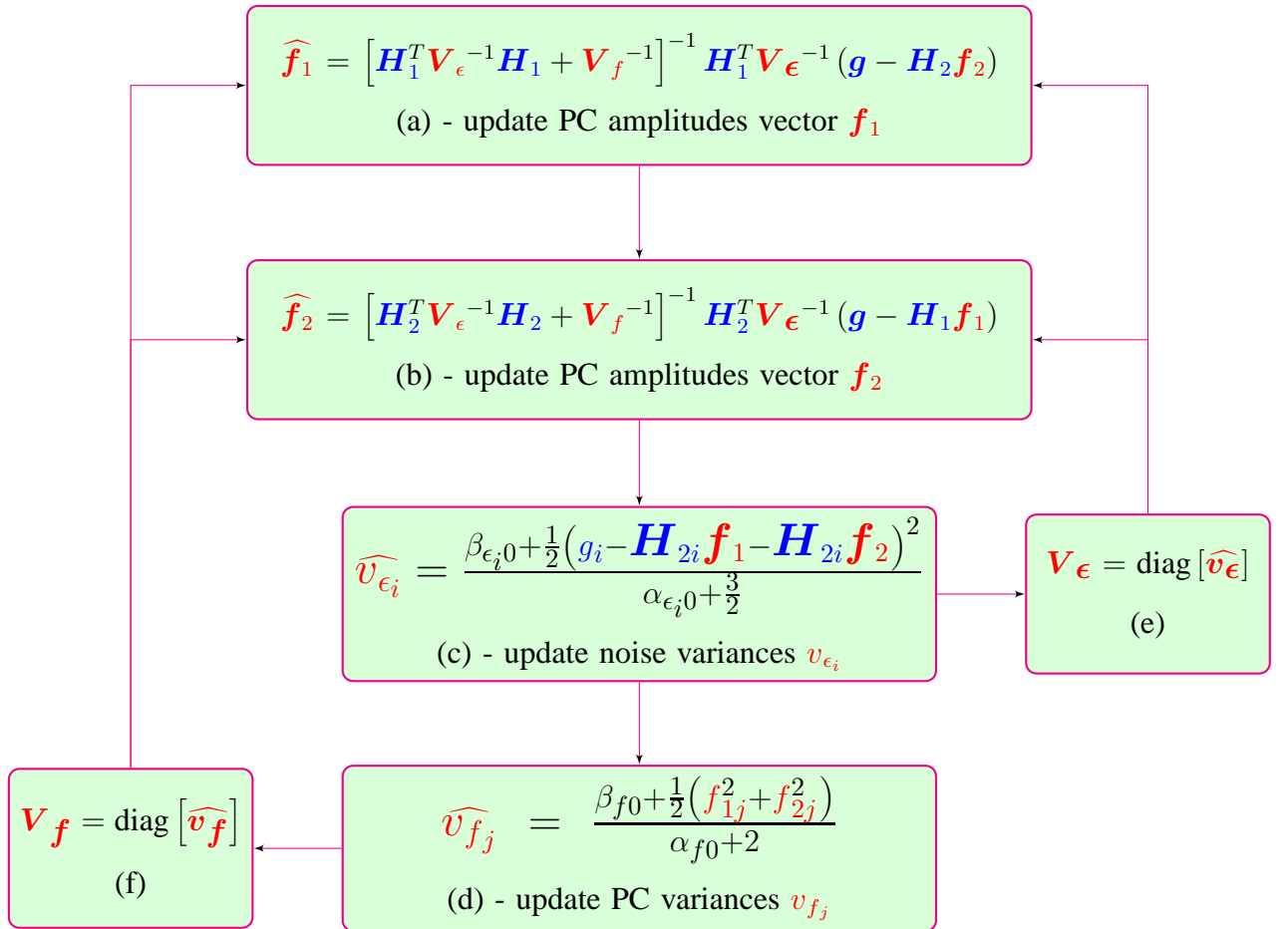


Figure 4.5: Updating scheme: JMAP estimation for generalized Student-t prior model

This algorithm is compared to the algorithm corresponding to Posterior Mean estimation in the simulations section.

#### 4.4.3 Posterior Mean (via VBA, partial separability)

The JMAP computes the mod of the posterior distribution. The Posterior Mean computes the mean of the posterior distribution. One of the advantages of this estimator is that it minimizes

#### 4.4.3 - Posterior Mean (via VBA, partial separability)

the Mean Square Error (MSE). Computing the posterior means of any unknown needs great dimensional integration. For example,

$$E_p \{ \mathbf{f}_1 \} = \iiint \mathbf{f}_1 p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) d\mathbf{f}_1 d\mathbf{f}_2 d\mathbf{v}_\epsilon d\mathbf{v}_f. \quad (4.35)$$

In general, these computations are not easy. One way to obtain approximate estimates is to approximate  $p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  by a separable one  $q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) = q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f | \mathbf{g})$ , then computing

$$E_q \{ \mathbf{f}_1 \} = \int \mathbf{f}_1 d\mathbf{f}_1 \quad (4.36)$$

If we can do this approximation intelligently, i.e. in such a way that  $E_q \{ \mathbf{f}_1 \}$  is the same as  $E_p \{ \mathbf{f}_1 \}$  we gain great amount of computational cost. In particular, for the proposed hierarchical model, Equation (4.31), the posterior distribution, Equation (4.32), 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  $p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  with a separable law  $q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$ :

$$p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \approx q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) = q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f) \quad (4.37)$$

by minimizing of the Kullback-Leibler divergence, defined as:

$$\begin{aligned} \text{KL}(q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) : p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})) &= \\ &= \iiint \dots \int q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \ln \frac{q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})}{p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})} d\mathbf{f}_1 d\mathbf{f}_2 d\mathbf{v}_\epsilon d\mathbf{v}_f \end{aligned} \quad (4.38)$$

where the following notations are used:

$$q_3(\mathbf{v}_\epsilon) = \prod_{i=1}^N q_{3i}(v_{\epsilon_i}) ; \quad q_4(\mathbf{v}_f) = \prod_{j=1}^M q_{4j}(v_{f_j}) ; \quad d\mathbf{v}_\epsilon = \prod_{i=1}^N dv_{\epsilon_i} ; \quad d\mathbf{v}_f = \prod_{j=1}^M dv_{f_j} \quad (4.39)$$

The minimization can be done via alternate optimization resulting the following proportionalities:

$$\left\{ \begin{array}{l} q_1(\mathbf{f}_1) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f)} \right\} \\ q_2(\mathbf{f}_2) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f)} \right\} \\ q_{3i}(v_{\epsilon_i}) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_{3-i}(v_{\epsilon_i}) q_4(\mathbf{v}_f)} \right\}, \quad i \in \{1, 2, \dots, N\} \\ q_{4j}(v_{f_j}) \propto \exp \left\{ \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_{4-j}(v_{f_j})} \right\}, \quad j \in \{1, 2, \dots, M\}, \end{array} \right. \quad (4.40)$$

using the notations:

$$q_{3-i}(v_{\epsilon_i}) = \prod_{k=1, k \neq i}^N q_{3k}(v_{\epsilon_k}) ; \quad q_{4-j}(v_{f_j}) = \prod_{k=1, k \neq j}^M q_{4k}(v_{f_k}) \quad (4.41)$$

and also

$$\langle u(x) \rangle_{v(y)} = \int u(x)v(y) dy. \quad (4.42)$$

From Equation (4.32), the analytical expression of the logarithm is obtained:

$$\begin{aligned}
 \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_{f_1}, \mathbf{v}_{f_2} | \mathbf{g}) &= C - \frac{1}{2} \ln \det(\mathbf{V}_\epsilon) - \frac{1}{2} \|\mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 \\
 &\quad - \sum_{i=1}^N (\alpha_{\epsilon_i 0} + 1) \ln v_{\epsilon_i} - \sum_{i=1}^N \beta_{\epsilon_i 0} v_{\epsilon_i}^{-1} - \ln \det(\mathbf{V}_f) \\
 &\quad - \frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1\|^2 - \sum_{j=1}^M (\alpha_{f_j 0} + 1) \ln v_{f_j} - \sum_{j=1}^M \beta_{f_j 0} v_{f_j}^{-1} \\
 &\quad - \frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_2\|^2
 \end{aligned} \tag{4.43}$$

• **Expression of  $q_1(\mathbf{f}_1)$  and  $q_2(\mathbf{f}_2)$ :**

The proportionality relation corresponding to  $q_1(\mathbf{f}_1)$  established in equation (4.40) refers to  $\mathbf{f}_1$ , so in the expression of  $\ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  all the terms free of  $\mathbf{f}_1$  can be regarded as constants:

$$\begin{aligned}
 \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f)} &= C \\
 &\quad - \left\langle \frac{1}{2} \|\mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 \right\rangle_{q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon)} \\
 &\quad - \left\langle \frac{1}{2} \|\mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1\|^2 \right\rangle_{q_4(\mathbf{v}_f)}
 \end{aligned}$$

Introducing the notations:

$$\begin{aligned}
 \widetilde{v}_{\epsilon_i} &= \langle v_{\epsilon_i}^{-1} \rangle_{q_{3i}(\mathbf{v}_{\epsilon_i})} ; \quad \widetilde{\mathbf{v}}_\epsilon = [\widetilde{v}_{\epsilon_1} \dots \widetilde{v}_{\epsilon_i} \dots \widetilde{v}_{\epsilon_N}]^T ; \quad \widetilde{\mathbf{V}}_\epsilon = \text{diag}[\widetilde{\mathbf{v}}_\epsilon] \\
 \widetilde{v}_{f_j} &= \langle v_{f_j}^{-1} \rangle_{q_{4j}(\mathbf{v}_{f_j})} ; \quad \widetilde{\mathbf{v}}_f = [\widetilde{v}_{f_1} \dots \widetilde{v}_{f_j} \dots \widetilde{v}_{f_M}]^T ; \quad \widetilde{\mathbf{V}}_f = \text{diag}[\widetilde{\mathbf{v}}_f]
 \end{aligned} \tag{4.44}$$

we can write:

$$\begin{aligned}
 \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f)} &= C - \frac{1}{2} \left\langle \|\widetilde{\mathbf{V}}_\epsilon^{\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 \right\rangle_{q_2(\mathbf{f}_2)} \\
 &\quad - \frac{1}{2} \|\widetilde{\mathbf{V}}_f^{\frac{1}{2}} \mathbf{f}_1\|^2
 \end{aligned}$$

Developing the norm  $\|\widetilde{\mathbf{V}}_\epsilon^{\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2$ , introducing the notation

$$\mathbf{f}_{2PM} = \langle \mathbf{f}_2 \rangle_{q_2(\mathbf{f}_2)} \tag{4.45}$$

and then excluding the  $\mathbf{f}_1$  free terms, we obtain:

$$\begin{aligned}
 \left\langle \|\widetilde{\mathbf{V}}_\epsilon^{\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2)\|^2 \right\rangle_{q_2(\mathbf{f}_2)} &= -2 \mathbf{g}^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_1 \mathbf{f}_1 + \mathbf{f}_1^T \mathbf{H}_1^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_1 \mathbf{f}_1 \\
 &\quad + 2 \mathbf{f}_{2PM}^T \mathbf{H}_2^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_1 \mathbf{f}_1
 \end{aligned} \tag{4.46}$$

so the proportionality for  $q_1(\mathbf{f}_1)$  can be expressed as:

$$q_1(\mathbf{f}_1) \propto \exp \left\{ -\frac{1}{2} \left( \mathbf{f}_1^T (\mathbf{H}_1^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_1 + \widetilde{\mathbf{V}}_f) \mathbf{f}_1 - 2 (\mathbf{g} - \mathbf{H}_2 \mathbf{f}_2)^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_1 \mathbf{f}_1 \right) \right\} \tag{4.47}$$

#### 4.4.3 - Posterior Mean (via VBA, partial separability)

i.e. the probability density function  $q_1(\mathbf{f}_1)$  is proportional to an exponential having a quadratic criterion, leading to the conclusion that  $q_1(\mathbf{f}_1)$  is a multivariate Normal distribution,  $q_1(\mathbf{f}_1) = \mathcal{N}(\mathbf{f}_1 | \mathbf{f}_{1PM}, \Sigma_{1PM})$ . Evidently, for determining the mean of the Normal distribution it is sufficient to minimize the quadratic criterion of the exponential. The covariance matrix is then obtained by identification, leading to:

$$q_1(\mathbf{f}_1) = \mathcal{N}(\mathbf{f}_1 | \mathbf{f}_{1PM}, \Sigma_{1PM}), \begin{cases} \mathbf{f}_{1PM} = (\mathbf{H}_1^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_1 + \widetilde{\mathbf{V}}_f)^{-1} \mathbf{H}_1^T \widetilde{\mathbf{V}}_\epsilon (\mathbf{g} - \mathbf{H}_2 \mathbf{f}_{2PM}) \\ \Sigma_{1PM} = (\mathbf{H}_1^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_1 + \widetilde{\mathbf{V}}_f)^{-1} \end{cases} \quad (4.48)$$

with  $\widetilde{\mathbf{V}}_\epsilon$ ,  $\widetilde{\mathbf{V}}_f$  and  $\mathbf{f}_{2PM}$  introduced in (4.44) and (4.45). For the expression  $q_2(\mathbf{f}_2)$  all the computations are the same and following the same developments, we conclude that  $q_2(\mathbf{f}_2)$  is a multivariate Normal distribution:

$$q_2(\mathbf{f}_2) = \mathcal{N}(\mathbf{f}_2 | \mathbf{f}_{2PM}, \Sigma_{2PM}), \begin{cases} \mathbf{f}_{2PM} = (\mathbf{H}_2^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_2 + \widetilde{\mathbf{V}}_f)^{-1} \mathbf{H}_2^T \widetilde{\mathbf{V}}_\epsilon (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_{1PM}) \\ \Sigma_{2PM} = (\mathbf{H}_2^T \widetilde{\mathbf{V}}_\epsilon \mathbf{H}_2 + \widetilde{\mathbf{V}}_f)^{-1} \end{cases} \quad (4.49)$$

#### • Expression of $q_{3i}(v_{\epsilon_i})$ :

Considering the terms free of  $v_{\epsilon_i}$  as constants, we have:

$$\begin{aligned} \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_{3-i}(v_{\epsilon_i}) q_4(\mathbf{v}_f)} &= C - \frac{1}{2} \ln v_{\epsilon_i} - (\alpha_{\epsilon 0} + 1) \ln v_{\epsilon_i} \\ &- \frac{1}{2} \left\langle \left\| \mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2) \right\|^2 \right\rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_{3-i}(v_{\epsilon_i})} - \beta_{\epsilon 0} v_{\epsilon_i}^{-1} \end{aligned} \quad (4.50)$$

For  $\left\langle \left\| \mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2) \right\|^2 \right\rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_{3-i}(v_{\epsilon_i})}$  we have the following development:

$$\begin{aligned} \left\langle \left\| \mathbf{V}_\epsilon^{-\frac{1}{2}} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_1 - \mathbf{H}_2 \mathbf{f}_2) \right\|^2 \right\rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_{3-i}(v_{\epsilon_i})} &= C \\ &+ v_{\epsilon_i}^{-1} \left( g_i - \mathbf{H}_1^i \mathbf{f}_{1PM} - \mathbf{H}_2^i \mathbf{f}_{2PM} \right)^2 \\ &+ v_{\epsilon_i}^{-1} \left( \mathbf{H}_1^i \Sigma_{1PM} \mathbf{H}_1^{iT} \right) \\ &+ v_{\epsilon_i}^{-1} \left( \mathbf{H}_2^i \Sigma_{2PM} \mathbf{H}_2^{iT} \right) \end{aligned} \quad (4.51)$$

so the proportionality relation corresponding to  $q_{3i}(v_{\epsilon_i})$ , Equation (4.40) becomes:

$$q_{3i}(v_{\epsilon_i}) \propto v_{\epsilon_i}^{-(\alpha_{\epsilon_i} + 1)} \exp \left\{ -\beta_{\epsilon_i} v_{\epsilon_i}^{-1} \right\} \quad (4.52)$$

So  $q_{3i}(v_{\epsilon_i})$  is an Inverse Gamma distribution with the following parameters:

$$q_{3i}(v_{\epsilon_i}) = \mathcal{IG}(v_{\epsilon_i} | \alpha_{\epsilon_i}, \beta_{\epsilon_i}), \begin{cases} \alpha_{\epsilon_i} = \alpha_{\epsilon 0} + \frac{1}{2} \\ \beta_{\epsilon_i} = \beta_{\epsilon 0} + \frac{1}{2} \left[ (g_i - \mathbf{H}_1^i \mathbf{f}_{1PM} - \mathbf{H}_2^i \mathbf{f}_{2PM})^2 + \right. \\ \left. (\mathbf{H}_1^i \Sigma_{1PM} \mathbf{H}_1^{iT} + \mathbf{H}_2^i \Sigma_{2PM} \mathbf{H}_2^{iT}) \right] \end{cases} \quad (4.53)$$

#### • Expression of $q_{4j}(v_{f_j})$ :

Considering the terms free of  $v_{f_j}$  as constants, we have:

$$\begin{aligned} \langle \ln p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \rangle_{q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_{4-j}(v_{f_j})} &= -\ln v_{f_j} - (\alpha_{f_0} + 1) \ln v_{f_j} \\ &- \frac{1}{2} \left\langle \left\| \mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1 \right\|^2 \right\rangle_{q_1(\mathbf{f}_1) q_{4-j}(v_{f_j})} - \frac{1}{2} \left\langle \left\| \mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_2 \right\|^2 \right\rangle_{q_2(\mathbf{f}_2) q_{4-j}(v_{f_j})} - \beta_{f_0} v_{f_j}^{-1} \end{aligned} \quad (4.54)$$

For  $\left\langle \left\| \mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1 \right\|^2 \right\rangle_{q_1(\mathbf{f}_1) q_{4-j}(v_{f_j})}$  and  $\left\langle \left\| \mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_2 \right\|^2 \right\rangle_{q_2(\mathbf{f}_2) q_{4-j}(v_{f_j})}$  we have the following relations:

$$\begin{cases} \left\langle \left\| \mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_1 \right\|^2 \right\rangle_{q_1(\mathbf{f}_1) q_{4-j}(v_{f_j})} = C + v_{f_j}^{-1} (f_{1jPM}^2 + \Sigma_{1jjPM}) \\ \left\langle \left\| \mathbf{V}_f^{-\frac{1}{2}} \mathbf{f}_2 \right\|^2 \right\rangle_{q_2(\mathbf{f}_2) q_{4-j}(v_{f_j})} = C + v_{f_j}^{-1} (f_{2jPM}^2 + \Sigma_{2jjPM}) \end{cases} \quad (4.55)$$

so the proportionality relation corresponding to  $q_{4j}(v_{f_j})$  established in equation (4.40) becomes:

$$q_{4j}(v_{f_j}) \propto v_{f_j}^{-(\alpha_{f_j}+1)} \exp \left\{ -\beta_{f_j} v_{f_j}^{-1} \right\} \quad (4.56)$$

We conclude that  $q_{4j}(v_{f_j})$  is an Inverse Gamma distribution with the following parameters:

$$q_{4j}(v_{f_j}) = \mathcal{IG} \left( v_{f_j} | \alpha_{f_j}, \beta_{f_j} \right), \begin{cases} \alpha_{f_j} = \alpha_{f_0} + 1 \\ \beta_{f_j} = \beta_{f_0} + \frac{1}{2} \left[ f_{1jPM}^2 + \Sigma_{1jjPM} + f_{2jPM}^2 + \Sigma_{2jjPM} \right] \end{cases} \quad (4.57)$$

The expressions of the parameters corresponding to the two multivariate Normal distributions  $q_1(\mathbf{f}_1)$  and  $q_2(\mathbf{f}_2)$  depend on  $\widetilde{\mathbf{V}}_\epsilon$  and  $\widetilde{\mathbf{V}}_f$ , introduced in (4.44). Since  $q_{3i}(v_{\epsilon_i})$  and  $q_{4j}(v_{f_j})$  are Inverse Gamma distributions it is easy to express the two matrices with the corresponding Inverse Gamma parameters:

$$\begin{aligned} \widetilde{v}_{\epsilon_i} &= \left\langle v_{\epsilon_i}^{-1} \right\rangle_{\mathcal{IG}(v_{\epsilon_i} | \alpha_{\epsilon_i}, \beta_{\epsilon_i})} = \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \Rightarrow \widetilde{\mathbf{V}}_\epsilon = \text{diag} \left[ \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \right] \\ \widetilde{v}_{f_j} &= \left\langle v_{f_j}^{-1} \right\rangle_{\mathcal{IG}(v_{f_j} | \alpha_{f_j}, \beta_{f_j})} = \frac{\alpha_{f_j}}{\beta_{f_j}} \Rightarrow \widetilde{\mathbf{V}}_f = \text{diag} \left[ \frac{\alpha_{f_j}}{\beta_{f_j}} \right] \end{aligned} \quad (4.58)$$

### • Algorithm:

Equations (4.48), (4.49), (4.53), (4.57), representing the distribution families of separable laws  $q$  and their corresponding parameters are leading to an iterative algorithm estimating all the parameters involved. The shape parameters corresponding to the Inverse Gamma distributions

#### 4.4.3 - Posterior Mean (via VBA, partial separability)

rest constants during the iterations. All the other parameters are updated, changing the value.

$$\begin{cases} \mathbf{f}_{1PM} = \left( \mathbf{H}_1^T \text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \\ \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \mathbf{H}_1 + \text{diag} \begin{bmatrix} \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \right)^{-1} \mathbf{H}_1^T \text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \end{bmatrix} (\mathbf{g} - \mathbf{H}_2 \mathbf{f}_{2PM}) \\ \Sigma_{1PM} = \left( \mathbf{H}_1^T \text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \\ \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \mathbf{H}_1 + \text{diag} \begin{bmatrix} \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \right)^{-1} \end{cases} \quad (4.59a)$$

$$\begin{cases} \mathbf{f}_{2PM} = \left( \mathbf{H}_2^T \text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \\ \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \mathbf{H}_2 + \text{diag} \begin{bmatrix} \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \right)^{-1} \mathbf{H}_2^T \text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \end{bmatrix} (\mathbf{g} - \mathbf{H}_1 \mathbf{f}_{1PM}) \\ \Sigma_{2PM} = \left( \mathbf{H}_2^T \text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \\ \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \mathbf{H}_2 + \text{diag} \begin{bmatrix} \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix} \right)^{-1} \end{cases} \quad (4.59b)$$

$$\begin{cases} \alpha_{\epsilon_i} = \alpha_{\epsilon 0} + \frac{1}{2} \\ \beta_{\epsilon_i} = \beta_{\epsilon 0} + \frac{1}{2} \left[ (g_i - \mathbf{H}_1^i \mathbf{f}_{1PM} - \mathbf{H}_2^i \mathbf{f}_{2PM})^2 + (\mathbf{H}_1^i \Sigma_{1PM} \mathbf{H}_1^{iT} + \mathbf{H}_2^i \Sigma_{2PM} \mathbf{H}_2^{iT}) \right] \end{cases} \quad (4.59c)$$

$$\begin{cases} \alpha_{f_j} = \alpha_{f 0} + 1 \\ \beta_{f_j} = \beta_{f 0} + \frac{1}{2} \left[ f_{1jPM}^2 + \Sigma_{1jjPM} + f_{2jPM}^2 + \Sigma_{2jjPM} \right] \end{cases} \quad (4.59d)$$

Equation (4.59a) establish the dependency of the parameters corresponding to the Multivariate Normal Distribution  $q_1(\mathbf{f}_1)$  and the others hyperparameters involved in the Hierarchical Model: the mean  $\mathbf{f}_{1PM}$  and the covariance matrix  $\Sigma_{1PM}$  depend on  $\text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \end{bmatrix}$  and  $\text{diag} \begin{bmatrix} \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix}$ :

$$\boxed{\{\beta_{\epsilon_i}\}, \{\beta_{f_j}\}} \longrightarrow \boxed{\mathbf{f}_{1PM}, \Sigma_{1PM}} \quad (4.60)$$

Equation (4.59b) establish the dependency of the parameters corresponding to the Multivariate Normal Distribution  $q_2(\mathbf{f}_2)$  and the others hyperparameters involved in the Hierarchical Model: the mean  $\mathbf{f}_{2PM}$  and the covariance matrix  $\Sigma_{2PM}$  depend on  $\text{diag} \begin{bmatrix} \frac{\alpha_{\epsilon_i}}{\beta_{\epsilon_i}} \end{bmatrix}$  and  $\text{diag} \begin{bmatrix} \frac{\alpha_{f_j}}{\beta_{f_j}} \end{bmatrix}$ :

$$\boxed{\{\beta_{\epsilon_i}\}, \{\beta_{f_j}\}} \longrightarrow \boxed{\mathbf{f}_{2PM}, \Sigma_{2PM}} \quad (4.61)$$

Equation (4.59c) leads to the following dependency scheme:

$$\boxed{\mathbf{f}_{1PM}, \Sigma_{1PM}, \mathbf{f}_{2PM}, \Sigma_{2PM}} \longrightarrow \boxed{\{\beta_{\epsilon_i}\}} \quad (4.62)$$

Equation (4.59d) leads to the following dependency schemes:

$$\boxed{\mathbf{f}_{1PM}, \Sigma_{1PM}, \mathbf{f}_{2PM}, \Sigma_{2PM}} \longrightarrow \boxed{\{\beta_{f_j}\}} \quad (4.63)$$

The scheme of the iterative algorithm is presented in Figure (4.6):

An important part of the algorithm is represented by the initialization part. We have to choose the parameters that will be used for the initialization part and assign them numerical values. We also have to assign numerical values for the shape and scale parameters corresponding to the two Inverse Gamma prior distributions involved in the hierarchical model. In

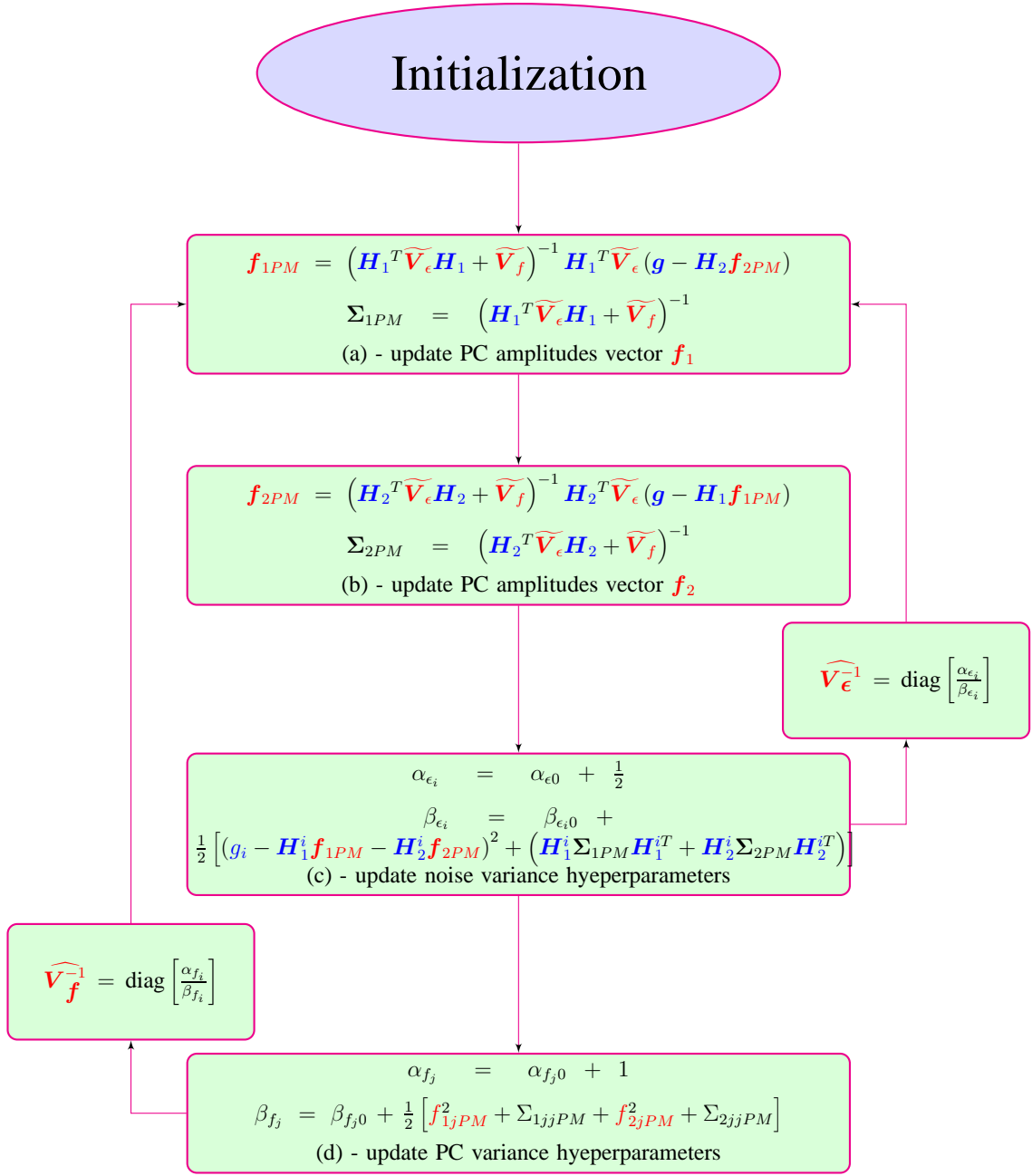


Figure 4.6: Updating scheme: PM estimation via VBA (partial separability) for generalized Student-t prior model

order to select the values corresponding to  $\alpha_{f_0}$ ,  $\beta_{f_0}$ ,  $\alpha_{\epsilon_0}$ ,  $\beta_{\epsilon_0}$  we have to account for the fact that those hyperparameters have a crucial role in the selection of the model. The marginal of the Normal-Inverse Gamma distribution, i.e. the generalized Student-t modelling the two PC amplitudes vectors  $\mathbf{f}_1$  and  $\mathbf{f}_2$  depends on  $\alpha_{f_0}$ ,  $\beta_{f_0}$ , Equation (A.25). Furthermore, for the mechanism sparsity,  $\alpha_{f_0}$ ,  $\beta_{f_0}$  must be chosen such as they impose a sparse structure for the variance vector  $\mathbf{v}_f$  via the Inverse Gamma distribution.  $\alpha_{\epsilon_0}$ ,  $\beta_{\epsilon_0}$  must be chosen such that the proposed algorithm can distinguish between the peaks from the PC amplitudes vectors corresponding to the biological phenomena and the peaks corresponding to the noise.



### 4.4.4 Prior Inverse Gamma parameters setting for $v_{f_j}$

The proposed hierarchical model (4.31) is modelling the variances corresponding to  $f_1$  and  $f_2$ , namely  $v_{f_j}$  as Inverse Gamma distributions.

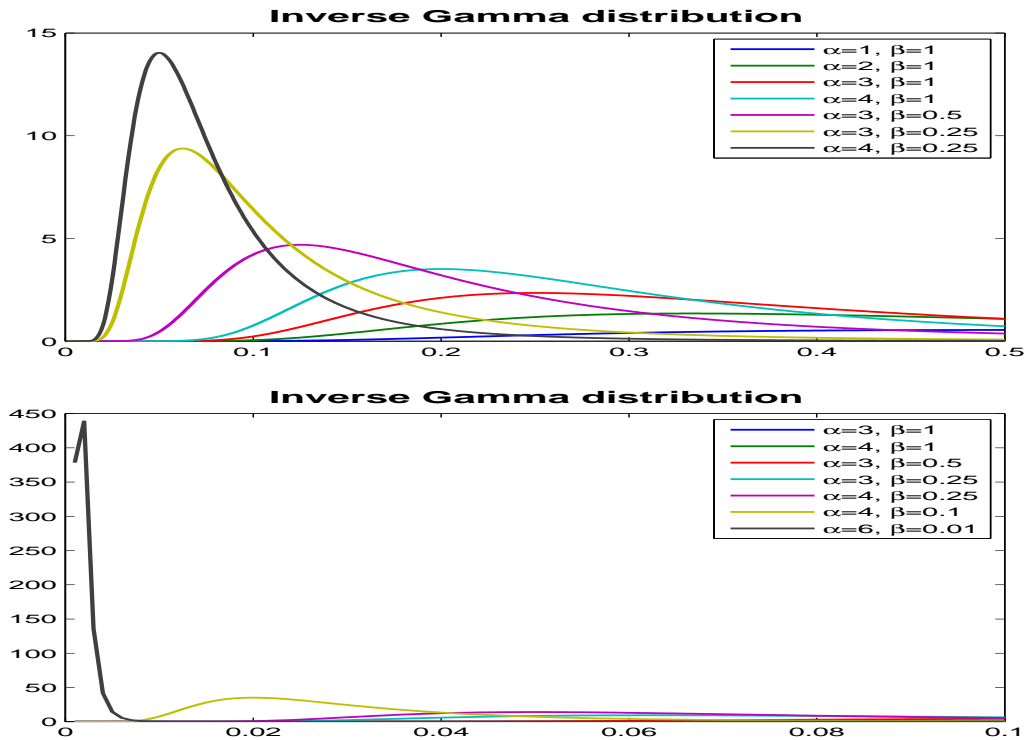


Figure 4.7: Inverse Gamma distributions for different  $\alpha, \beta$  parameters

Figure (4.7) presents the behaviour of the Inverse Gamma distribution, depending on the two parameters. The fundamental goal of the proposed hierarchical model was to search a sparse solution for  $f_1$  and  $f_2$ . The mechanism of the sparsity enforcing is linked with the

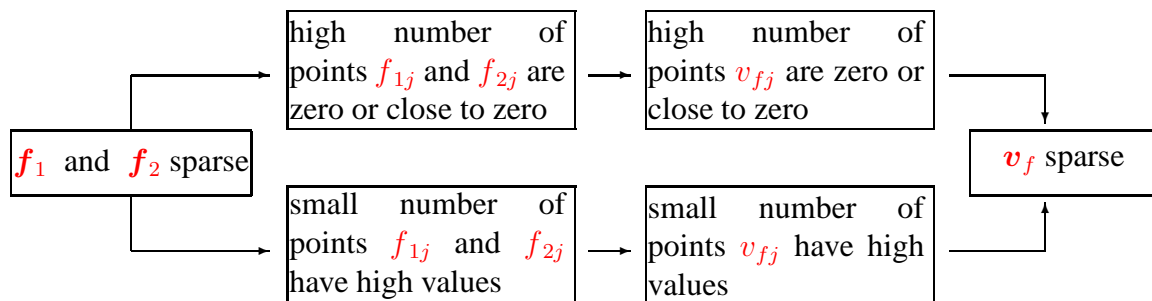


Figure 4.8: Sparsity mechanism: the sparsity is imposed (also) via the corresponding variance

corresponding variances  $v_{f_j}$ : the points of the vector with zero or almost zero values are associated with small variances, meaning that once the algorithm is setting those values close to zero then they will rest close to zero, not having high variations. For the non-zero values, we want to allow certain variations of the amplitude. In this case, the variances associated

must have significant values. The link between the PC amplitudes vectors  $\mathbf{f}_1$  and  $\mathbf{f}_2$  and the variances vector  $\mathbf{v}_f$  is presented in Figure (4.8). Therefore, we search the Inverse Gamma distributions parameters such that the expected value is close to zero and the variance is close to a certain value  $w$  set in accordance with the data.

The expected value of an Inverse Gamma distribution,  $\mathcal{IG}(x|\alpha, \beta)$  is defined for  $\alpha > 1$ ,  $E_{\mathcal{IG}}[x] = \frac{\beta}{\alpha-1}$ , and the variance is defined for  $\alpha > 2$ ,  $Var_{\mathcal{IG}}[x] = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)}$ . The shape and scale parameters for the Inverse Gamma distribution must be set such that:

$$\frac{\beta}{\alpha-1} = \epsilon \simeq 0 \quad ; \quad \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} = w \quad (4.64)$$

One way to assure an Inverse Gamma distribution that presents a small expected values  $\epsilon \simeq 0$  and a certain  $w$  variation is to adopt the following model for the parameters:

$$\alpha = 2 + \epsilon_0^2 \quad ; \quad \beta = \sqrt{w}\epsilon_0(1 + \epsilon_0^2) \quad , \quad \epsilon_0 = \frac{\epsilon}{\sqrt{w}} \quad (4.65)$$

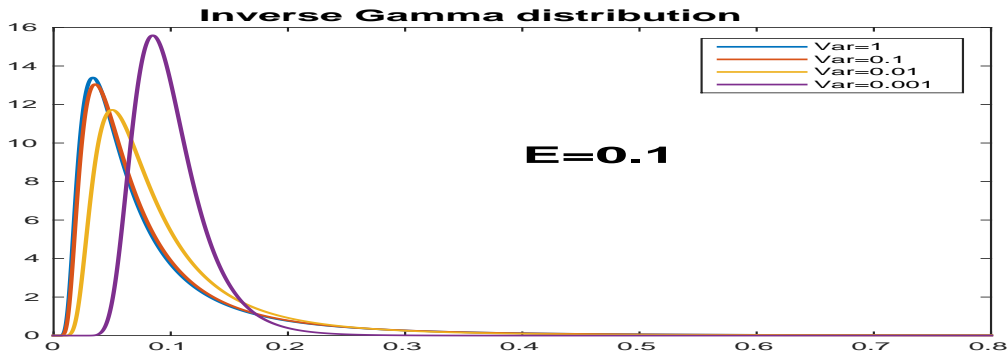


Figure 4.9:  $x \sim \mathcal{IG}(x|\alpha, \beta)$ ;  $E[x] = 0.1$ ;  $Var[x] = 1, 0.1, 0.01$  and  $0.001$

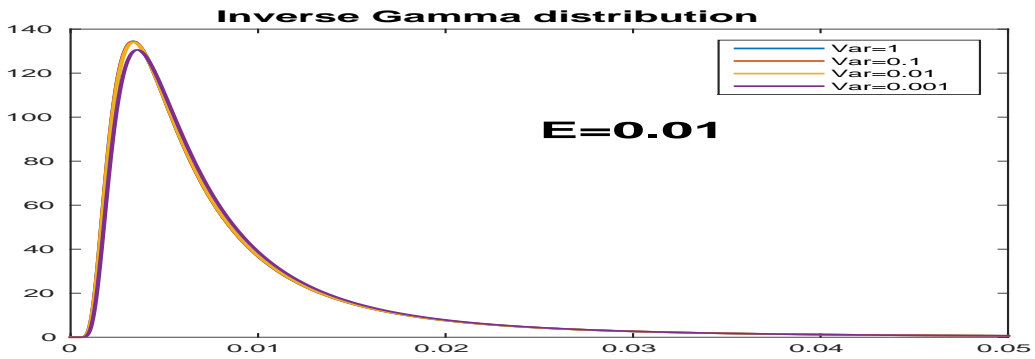


Figure 4.10:  $x \sim \mathcal{IG}(x|\alpha, \beta)$ ;  $E[x] = 0.01$ ;  $Var[x] = 1, 0.1, 0.01$  and  $0.001$

Figure (4.9) shows four Inverse Gamma distributions for which the mean is the same,  $E[x] = 0.1$ , but with the variances  $Var[x] = 1, 0.1, 0.01$  and  $0.001$ , for which the parameters are computed using Equation (4.65). The generalized Student-t distributions, i.e. the marginals of the IGSM distributions corresponding to the two parameters  $\alpha$  and  $\beta$ , imposing the mean and variances mentioned above are presented in Figures (4.12), (4.13)

#### 4.4.4 - Prior Inverse Gamma parameters setting for $v_{f_j}$

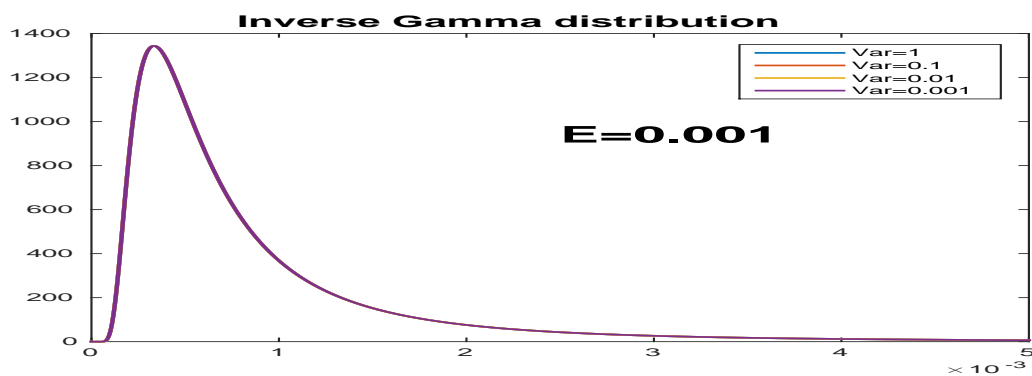


Figure 4.11:  $x \sim \mathcal{IG}(x|\alpha, \beta)$ ;  $E[x] = 0.001$ ;  $\text{Var}[x] = 1, 0.1, 0.01$  and  $0.001$

$\epsilon$	$w$	$\text{Var}_{St_g} = E_{\mathcal{IG}}$	$\text{Var}_{\mathcal{IG}}$	Figure
0.1	0.1	0.1	0.1	Figures (4.9), (4.12),(4.13)
0.1	0.01	0.1	0.01	Figures (4.9),(4.12),(4.13)
0.1	0.001	0.1	0.001	Figures (4.9),(4.12),(4.13)
0.01	0.1	0.01	0.1	Figure (4.10),(4.14),(4.15)
0.01	0.01	0.01	0.01	Figure (4.10),(4.14),(4.15)
0.01	0.001	0.01	0.001	Figure (4.10),(4.14),(4.15)
0.001	0.1	0.001	0.1	Figure (4.11),(4.16),(4.17)
0.001	0.01	0.001	0.01	Figure (4.11),(4.16),(4.17)
0.001	0.001	0.001	0.001	Figure (4.11),(4.16),(4.17)

Table 4.1: Mean and variance for  $St_g$  via  $\epsilon$  and  $w$

Figure (4.10) shows four Inverse Gamma distributions for which the mean is the same,  $E[x] = 0.01$ , but with the variances  $\text{Var}[x] = 1, 0.1, 0.01$  and  $0.001$ , for which the parameters are computed using Equation (4.65). The generalized Student-t distributions, i.e. the marginals of the IGSM distributions corresponding to the two parameters  $\alpha$  and  $\beta$ , imposing the mean and variances mentioned above are presented in Figures (4.14), (4.15)

Figure (4.11) shows four Inverse Gamma distributions for which the mean is the same,  $E[x] = 0.001$ , but with the variances  $\text{Var}[x] = 1, 0.1, 0.01$  and  $0.001$ , for which the parameters are computed using Equation (4.65). The generalized Student-t distributions, i.e. the marginals of the IGSM distributions corresponding to the two parameters  $\alpha$  and  $\beta$ , imposing the mean and variances mentioned above are presented in Figures (4.16), (4.17)

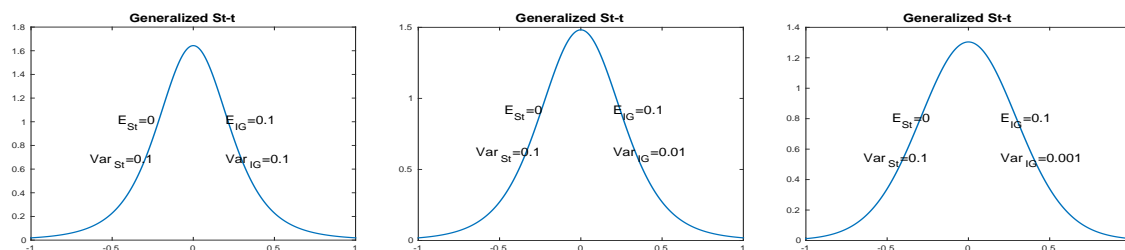


Figure 4.12: Different Generalized Student-t distributions for which the means  $E_{St_g} = 0$  and the variances  $\text{Var}_{St_g} = 0.1$  are the same

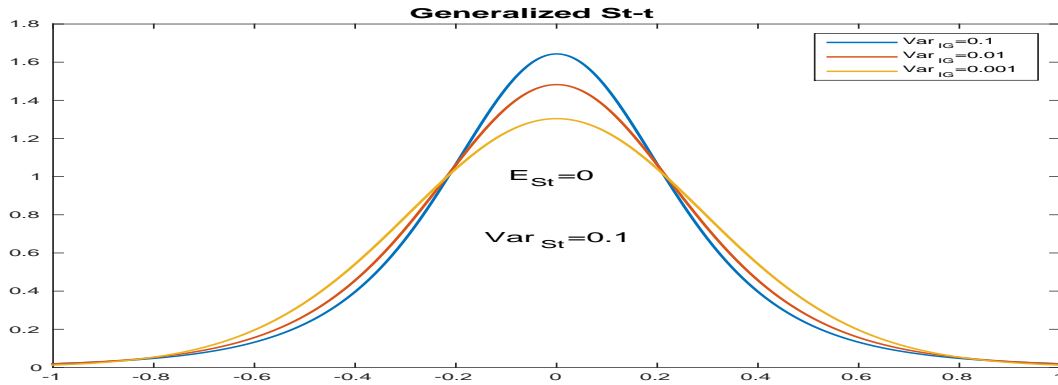


Figure 4.13: Generalized Student-t:  $\text{Var}_{St_g} = 0.1$ ;  $\text{Var}_{IG} = 0.1, 0.01, 0.001$ ;

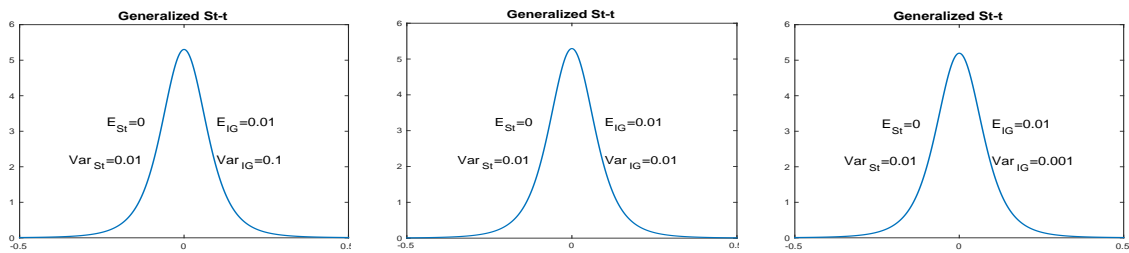


Figure 4.14: Different Generalized Student-t distributions for which the means  $E_{St_g} = 0$  and the variances  $\text{Var}_{St_g} = 0.01$  are the same

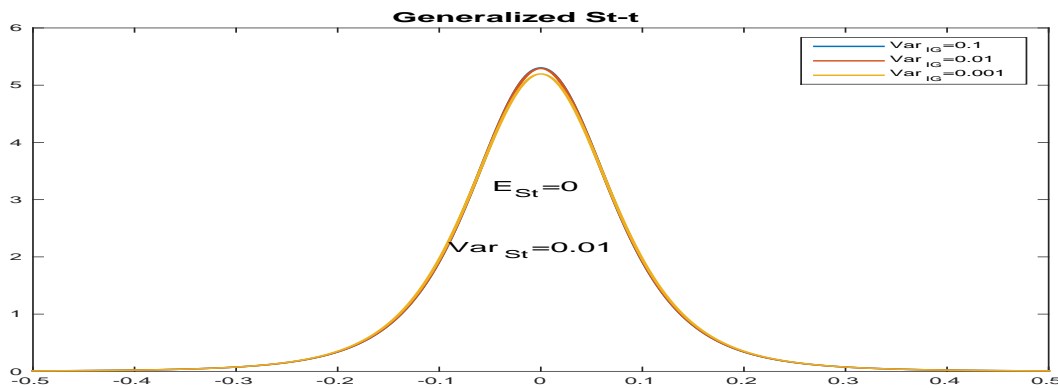


Figure 4.15: Generalized Student-t:  $\text{Var}_{St_g} = 0.01$ ;  $\text{Var}_{IG} = 0.1, 0.01, 0.001$ ;

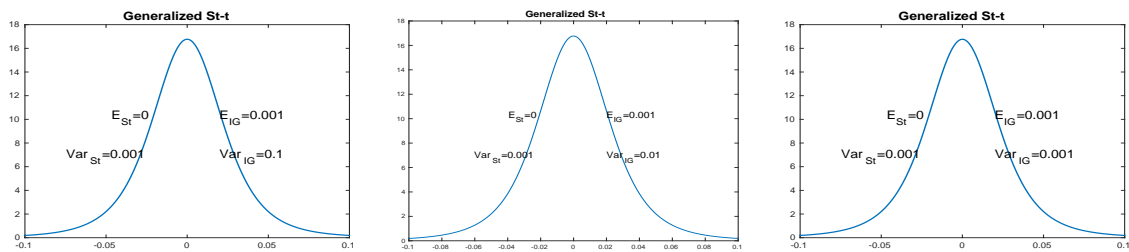


Figure 4.16: Different Generalized Student-t distributions for which the means  $E_{St_g} = 0$  and the variances  $\text{Var}_{St_g} = 0.001$  are the same

#### 4.4.5 Prior Inverse Gamma parameters setting for $v_{\epsilon_i}$

The variance corresponding to the noise,  $v_{\epsilon_i}, i \in 1, 2, \dots, N$  are also modelled as Inverse Gamma distributions, for the same conjugacy reasons as in the case of  $v_{f_j}$ . However, while

#### 4.4.5 - Prior Inverse Gamma parameters setting for $v_{\epsilon_i}$

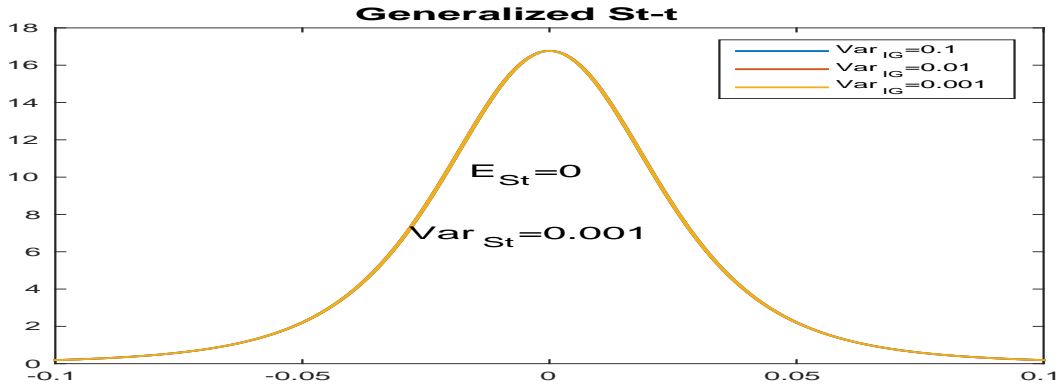


Figure 4.17: Generalized Student-t:  $\text{Var}_{St_g} = 0.001$ ;  $\text{Var}_{\mathcal{IG}} = 0.1, 0.01, 0.001$ ;

the behaviour of  $v_{f_j}$  variances is linked with the sparsity enforcing of  $f_1$  and  $f_2$ , so the parameters of the prior Inverse Gamma distributions  $\mathcal{IG}(v_{f_j} | \alpha_{f_0}, \beta_{f_0})$  and can be fixed such that the mean and the variance are in accordance with the sparsity enforcing strategy, it is not the case for the noise variance. Therefore, in order to fix the parameters of the prior Inverse Gamma distributions corresponding to the noise variance, another strategy is needed. We propose two possible approaches.

#### Non-Informative Prior law for $v_{\epsilon_i}$

As mentioned above, for fixing the parameters of Inverse Gamma distributions for  $v_{f_j}$  the prior knowledge was used. In the case of the noise variance, no (direct) prior information is available. Therefore, we can consider that the most proper way to model the noise variance is a non-informative prior law. So, in this case, the Inverse Gamma distribution parameters must be set such that the distribution becomes non-informative, i.e. uniform. The Inverse Gamma distribution becomes uniform for the parameters equal to zero:

$$\alpha_{\epsilon_0} = 0 ; \beta_{\epsilon_0} = 0 \quad (4.66)$$

#### Data deriving for $v_{\epsilon_i}$

The second approach for fixing the Inverse Gamma distribution parameters corresponding to  $v_{\epsilon_i}$  such that the mean and the variance correspond to numerical values derived from data. In a pure Bayesian inference, such approach could be criticized, since doing this, the available data, namely the known chronobiological signal  $g$ , is used twice for inference: first when deriving the numerical values corresponding to the mean and variance of the prior Inverse Gamma distribution modelling the noise variance and secondly during the proposed algorithm. However, for the application addressed in this thesis the main limitations is the lack of prior information. Therefore, this we consider also this approach. The mean and the variance of  $v_{\epsilon_i}$  are expressed via the two Inverse Gamma parameters,  $\alpha_{\epsilon_0}$  and  $\beta_{\epsilon_0}$ :

$$E[v_{\epsilon_i}] = \frac{\beta_{\epsilon_0}}{\alpha_{\epsilon_0} - 1}, \text{ for } \alpha_{\epsilon_0} > 1 \quad , \quad \text{Var}[v_{\epsilon_i}] = \frac{\beta_{\epsilon_0}^2}{(\alpha_{\epsilon_0} - 1)^2 (\alpha_{\epsilon_0} - 2)}, \text{ for } \alpha_{\epsilon_0} > 2 \quad (4.67)$$

Both values,  $E[v_{\epsilon_i}]$  and  $\text{Var}[v_{\epsilon_i}]$  correspond to the noise variance (the mean of the noise variance and the variance of the noise variance), an unknown of our model. The known input

$\mathbf{g}$  contains the error vector. This information can be exploited to obtain numerical values for the Inverse Gamma distribution mean and variance. Considering the mean of the known data  $E[\mathbf{g}]$ , the mean of the noise variance  $E[v_\epsilon]$  can be defined as the squared norm of the difference between the data and its mean, while the variance of the noise variance  $\text{Var}[v_\epsilon]$  can be defined in the same manner:  $E[v_\epsilon] = \|\mathbf{g} - E[\mathbf{g}]\|^2$ ,  $\text{Var}[v_\epsilon] = \|(\mathbf{g} - E[\mathbf{g}]) - E[\mathbf{g} - E[\mathbf{g}]]\|^2$ . Equations (4.67) links the parameters  $\alpha_{\epsilon_0}$  and  $\beta_{\epsilon_0}$  with numerical values that are derived from the known data:

$$\alpha_{\epsilon_0} = \frac{E[v_\epsilon]}{\text{Var}[v_\epsilon]} + 2 ; \quad \beta_{\epsilon_0} = E[v_\epsilon] \left( \frac{E[v_\epsilon]}{\text{Var}[v_\epsilon]} + 1 \right) \quad (4.68)$$

In particular, such an approach, is consistent to a non-supervised algorithm.

#### 4.4.6 Posterior Mean (via VBA, full separability)

In Subsection (4.4.3) the Student-t model is considered and the PM estimator is used. The posterior law is approximated by a separable one, Equation (4.37), where the notations for  $q_3(\mathbf{v}_\epsilon)$  and  $q_4(\mathbf{v}_f)$ , introduced in Equation (4.39), represent a full separability relative to  $\mathbf{v}_\epsilon$  and  $\mathbf{v}_f$ . But the prior law  $p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  is not approximated by a fully separable one since for  $q_1(\mathbf{f}_1)$  and  $q_2(\mathbf{f}_2)$  we consider a multivariate law modelling the vector  $\mathbf{f}_1$  and  $\mathbf{f}_2$ . In this subsection we investigate the development of the proposed model and 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:

$$p(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) \approx q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g}) = q_1(\mathbf{f}_1) q_2(\mathbf{f}_2) q_3(\mathbf{v}_\epsilon) q_4(\mathbf{v}_f) \quad (4.69)$$

where the notations for  $q_3(\mathbf{v}_\epsilon)$  and  $q_4(\mathbf{v}_f)$  are defined in Equation (4.39) and :

$$q_1(\mathbf{f}_1) = \prod_{j=1}^M q_{1j}(f_{1j}) ; \quad q_2(\mathbf{f}_2) = \prod_{j=1}^M q_{2j}(f_{2j}) ; \quad d\mathbf{f}_1 = \prod_{j=1}^M df_{1j} ; \quad d\mathbf{f}_2 = \prod_{j=1}^M df_{2j} \quad (4.70)$$

and

$$q_{1-j}(f_{1j}) = \prod_{k=1, k \neq j}^M q_{1k}(f_{1k}) ; \quad q_{2-j}(f_{2j}) = \prod_{k=1, k \neq j}^M q_{2k}(f_{2k}) \quad (4.71)$$

Like in Subsection (4.4.3), the law  $q(\mathbf{f}_1, \mathbf{f}_2, \mathbf{v}_\epsilon, \mathbf{v}_f | \mathbf{g})$  is obtained by minimizing the Kullback-Leibler divergence, Equation (4.38), via alternate optimization, obtaining an iterative algorithm. The scheme of the iterative algorithm is presented in Figure (4.18):

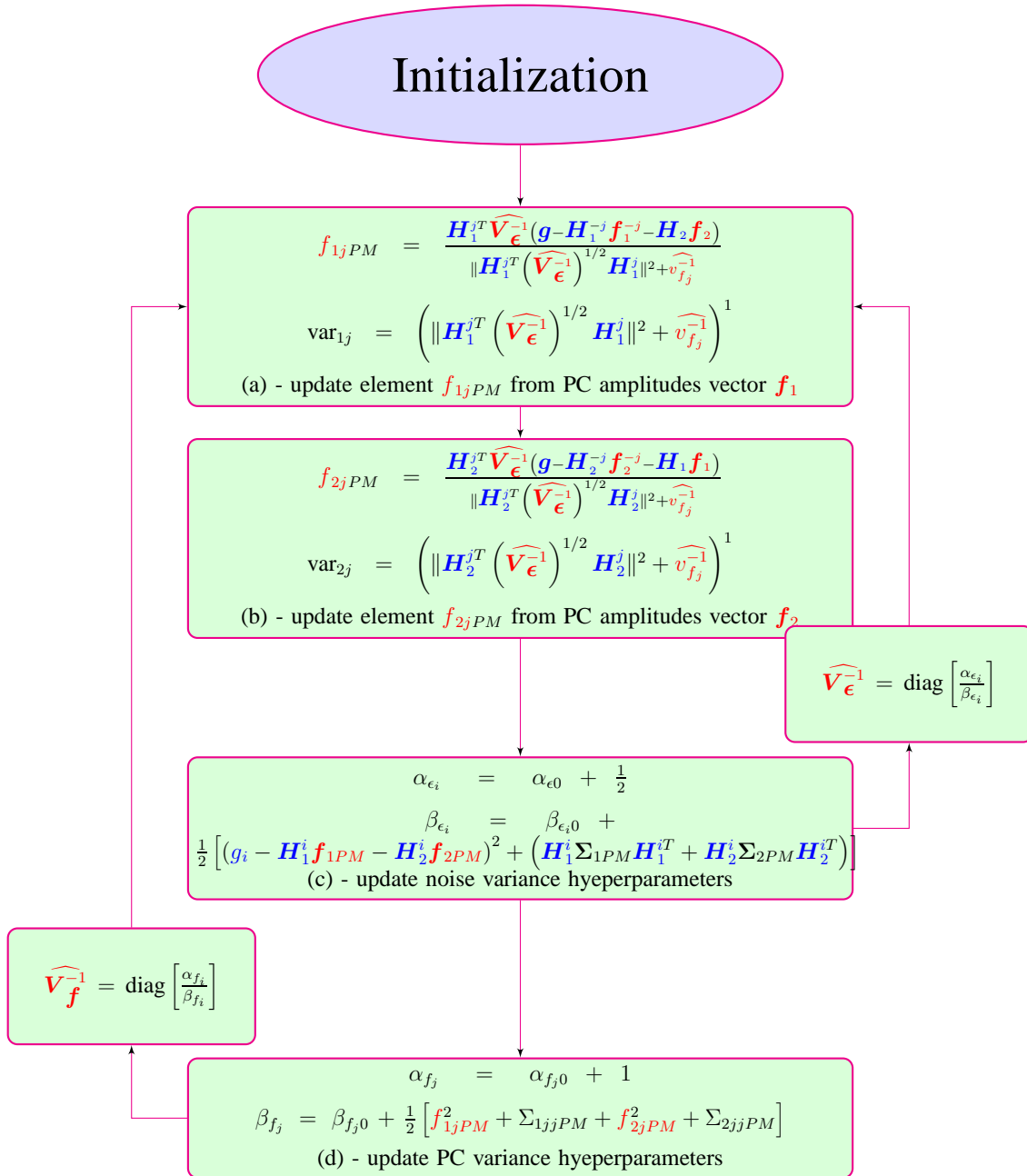


Figure 4.18: Updating scheme: PM estimation via VBA (full separability) for generalized Student-t prior model

# 5

## Simulations: Synthetic data

### Contents

---

<b>5.1 Simulations 05dB</b>	<b>80</b>
5.1.1 Synthetic data 05dB	80
5.1.2 Data 05dB	81
5.1.3 JMAP IGSM 05dB	82
5.1.4 PM (via VBA, partial separability) IGSM 05dB	84
5.1.5 PM (via VBA, full separability) IGSM 05dB	85
5.1.6 Methods comparison 05dB	86
5.1.7 Error comparison 05dB	87
<b>5.2 Synthetic data 10dB</b>	<b>91</b>
5.2.1 Data 10dB	91
5.2.2 JMAP IGSM 10dB	92
5.2.3 PM (via VBA, partial separability) IGSM 10dB	93
5.2.4 PM (via VBA, full separability) IGSM 10dB	95
5.2.5 Methods comparison 10dB	96
5.2.6 Error comparison 10dB	97
<b>5.3 Synthetic data 15dB</b>	<b>100</b>
5.3.1 Data 15dB	100
5.3.2 JMAP IGSM 15dB	101
5.3.3 PM (via VBA, partial separability) IGSM 15dB	102
5.3.4 PM (via VBA, full separability) IGSM 15dB	104
5.3.5 Methods comparison 15dB	105
5.3.6 Error comparison 15dB	106



### 5.1.1 - Synthetic data 05dB

---

<b>5.4 Student-t PM via VBA vs. LASSO 05dB</b> . . . . .	<b>109</b>
<b>5.5 Student-t PM via VBA vs. LASSO 10dB</b> . . . . .	<b>114</b>

---

THIS chapter is dedicated to the simulations results corresponding to the synthetic data, for the iterative algorithms presented in Chapter (4). First, we present in detail the results for three levels of noise, 05dB, Section (5.1), 10dB, Section (5.2) and 15dB, Section (5.3), for comparing the Gaussian model (i.e the Gaussian prior) with the proposed Student-t model. For each prior, we compare the algorithms corresponding to the JMAP and PM estimation. For all three sections, we present comparisons between the theoretical PC amplitude vector  $\mathbf{f}$  and the estimated PC amplitude vector  $\hat{\mathbf{f}}$ . We show the convergency of the PC vectors during the iterations and show the evolutions of the estimation errors for different realisation of the noise. Then, in Section (5.4) and Section (5.5) and we present the comparison between the Proposed Method and LASSO method, showing in detail the results corresponding to SNR=05dB and SNR=10dB and finally showing the behaviour of the  $L_1$  and  $L_2$  error norms corresponding to the Proposed Method and LASSO for SNR={50dB,40dB,30dB,25dB,20dB,15dB,10dB,05dB}.

For Sections (5.1), (5.2) and (5.3), five algorithms corresponding the Bayesian approach are compared: Joint MAP with Gaussian prior, Posterior Mean with Gaussian prior, Joint MAP with Student-t prior, Posterior Mean (via VBA) with Student-t prior (partial separability) and Posterior Mean (via VBA) with Student-t 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  $\hat{\mathbf{f}}_{Method}$  and  $\mathbf{f}$ , between  $\hat{\mathbf{g}}_{Method}$  and  $\mathbf{g}$  and between  $\hat{\mathbf{g}}_{Method}$  and  $\mathbf{g}_0$  theoretical signal ( $\mathbf{g}$  without noise). For every algorithm considered, we present the convergency 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 realisation, 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. The comparisons between error vectors corresponding to each algorithm are presented at the end of the subsections.

We consider the following protocol: we consider a theoretical PC vector  $\mathbf{f}$  and the corresponding theoretical signal  $\mathbf{H}\mathbf{f}$ ; we consider the corresponding signal  $\mathbf{g} = \mathbf{H}\mathbf{f} + \epsilon$ , by adding noise over the theoretical signal. The considered signal represents a four day signal, sampled every hour. The matrix  $\mathbf{H}$  considered in this set of simulations is a cosine plus sine matrix.

## 5.1 Simulations 05dB

---

This section presents the simulations corresponding to the synthetic case, for 05dB.

### 5.1.1 Synthetic data 05dB

For testing we have considered a four days signal, corresponding to a sparse PC vector, having non-zero values for 11h, 15h, and 23h. We consider this particular structure for the following

reasons: 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 an important peak corresponding to 24h and the corresponding harmonics, but this method offers no information for certain values in the circadian domain. We have showed in Chapter (3), Figure (3.2) that a dominant period, corresponding to 23h is wrongly estimated at 24h via FFT method and offers no other informations for the interval [20-31].

### 5.1.2 Data 05dB

The PC vector  $f$ , theoretical signal  $g_0$  and the signal  $g$  are presented in the Figure (5.1): Figure (5.1a) shows the theoretical PC, having the non-zero periods corresponding to 11h,

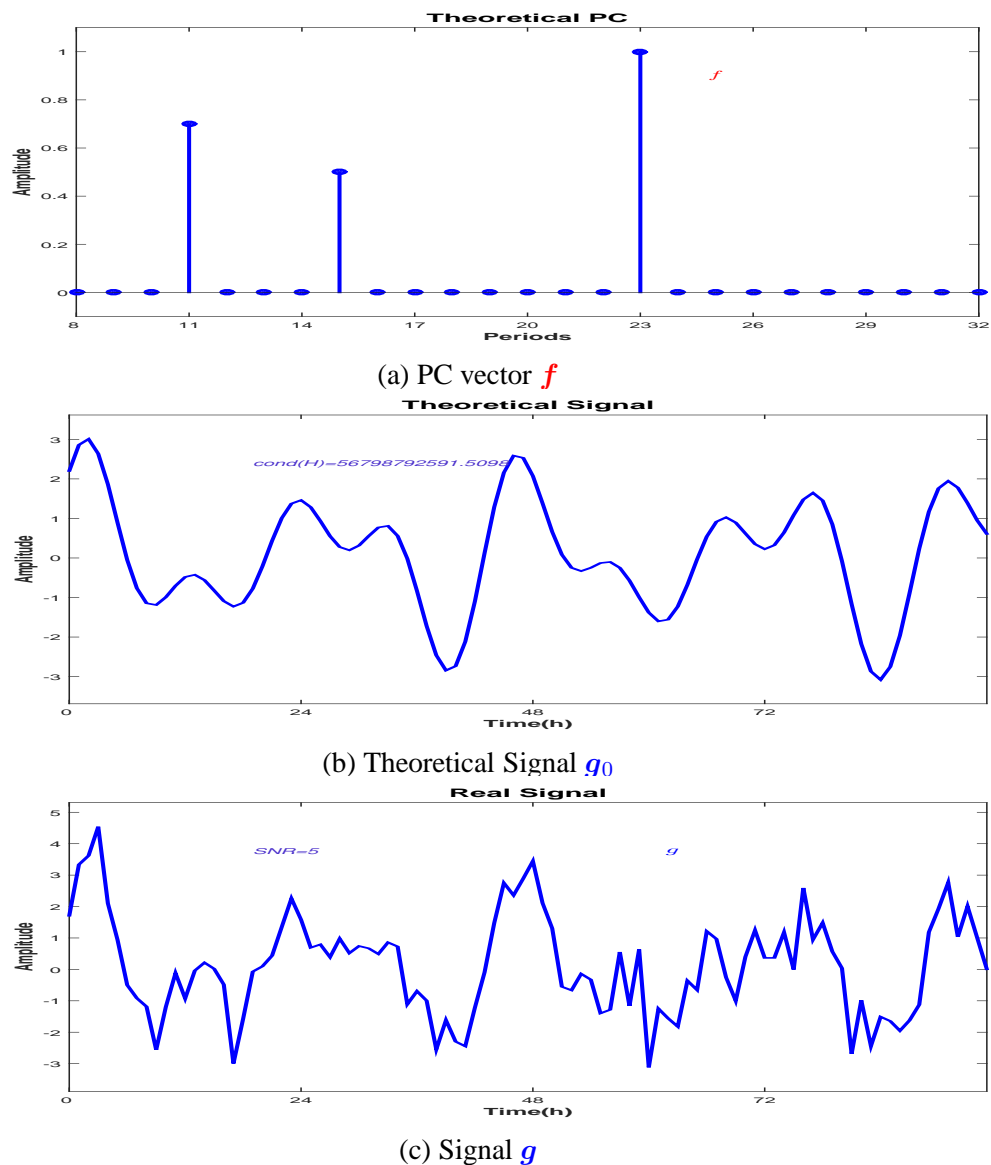


Figure 5.1:  $f$  PC vector, theoretical signal  $g_0$  and input signal  $g = g_0 + \epsilon$  of the model (5dB)

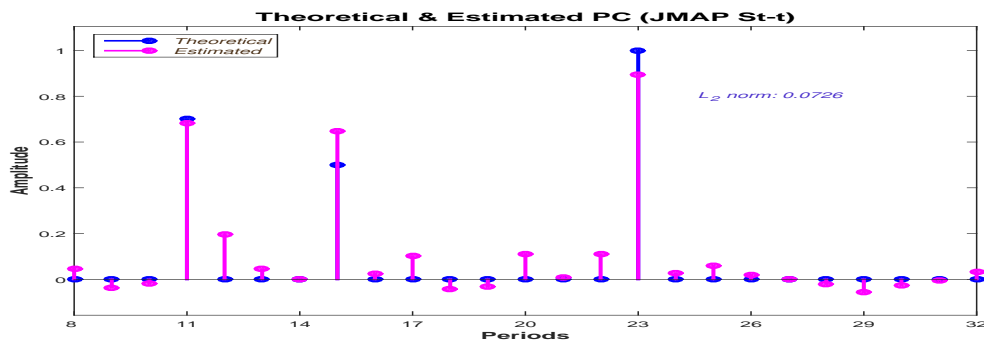
15h and 23h. All the other values in the PC vector are zero. Figure (5.1b) presents the signal corresponding to the linear model considered in Equation (4.4), neglecting the errors,  $g_0 = Hf$ . We note that the conditioning number of the matrix  $H$  is  $cond(H) = 56798792591$ .

### 5.1.3 - JMAP IGSM 05dB

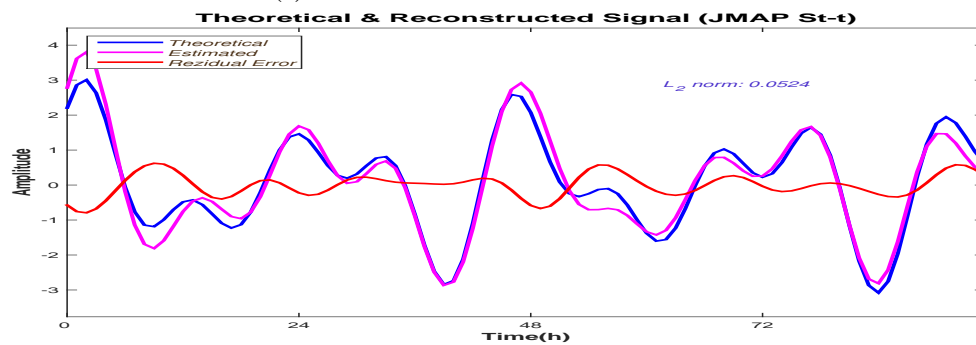
All the simulations are done using as the input the noisy signal  $g$  corresponding to the linear model, Equation (4.4), presented in Figure (5.1c). We compare the estimated PC vector with the theoretical one, Figure (5.1a) and the corresponding reconstructed signal with  $g_0$  and  $g$ . The comparison with the theoretical signal  $g_0$  is important in order to verify if the propose algorithm can distinguish the peaks corresponding to biological phenomena from the ones corresponding to the noise.

### 5.1.3 JMAP IGSM 05dB

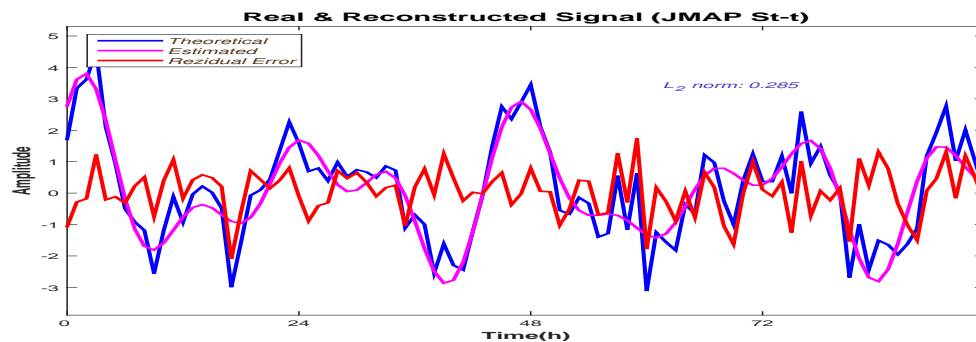
A comparison between the synthetic data and the JMAP estimation, corresponding to the IGSM prior hierarchical model is presented in Figure (5.2). We compare the theoretical PC vector  $f$  and the JMAP estimation  $\hat{f}_{JMAP}$ . We also present the comparison between the estimated  $\hat{g}_{JMAP}$  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 correspond-



(a) Theoretical and Estimated PC



(b) Theoretical and Estimated Signal



(c) Real and Estimated Signal

Figure 5.2: JMAP IGSM Estimation (5dB)

ing to the linear model, Equation (4.4). The comparison between the theoretical signal  $\mathbf{g}_0$  and  $\hat{\mathbf{g}}_{JMAP}$ , Figure (5.2b) shows that the proposed algorithm is converging to a solution that leads to a fairly accurate reconstruction, having the  $L_2$  norm error  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{JMAP}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.0524$ .

For the PC vector, the reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{JMAP}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.0726$ . For the JMAP estimation, the condition imposed for the searched solution, i.e. the sparsity is not respected, Figure (5.2a). 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 (5.3a) presents the vari-

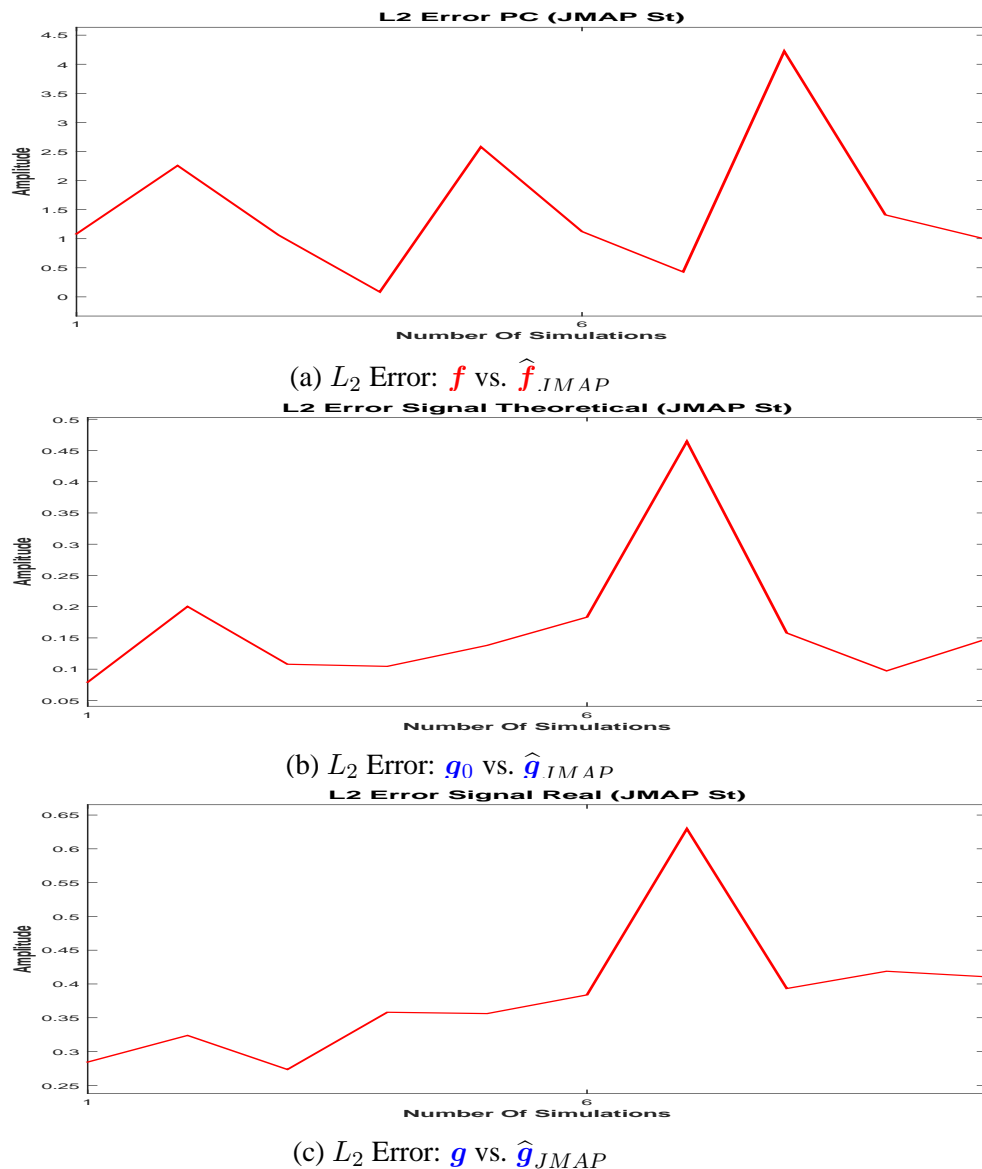


Figure 5.3: JMAP IGSM  $L_2$  error measured for 10 different noise realisations (5dB)

ation of  $L_2$  PC vector error reconstruction for 10 different noise realisation. 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  $\mathbf{g}_0$ , and signal  $\mathbf{g}$ , are presented in Figure (5.3b) and Figure (5.3c).

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

A comparison between the synthetic data and the PM (via VBA, partial separability) IGSM estimation is presented in Figure (5.4). We compare the theoretical PC vector  $\mathbf{f}$  with the PM (via VBA, partial separability) IGSM estimation  $\hat{\mathbf{f}}_{PM}$ , Figure (5.4a) and the corresponding reconstructed signal  $\hat{\mathbf{g}}_{PM}$  both with the theoretical signal  $\mathbf{g}_0$ , Figure (5.4b) and the input signal  $\mathbf{g}$ , Figure (5.4c). In the case of the Posterior Mean estimation via VBA,

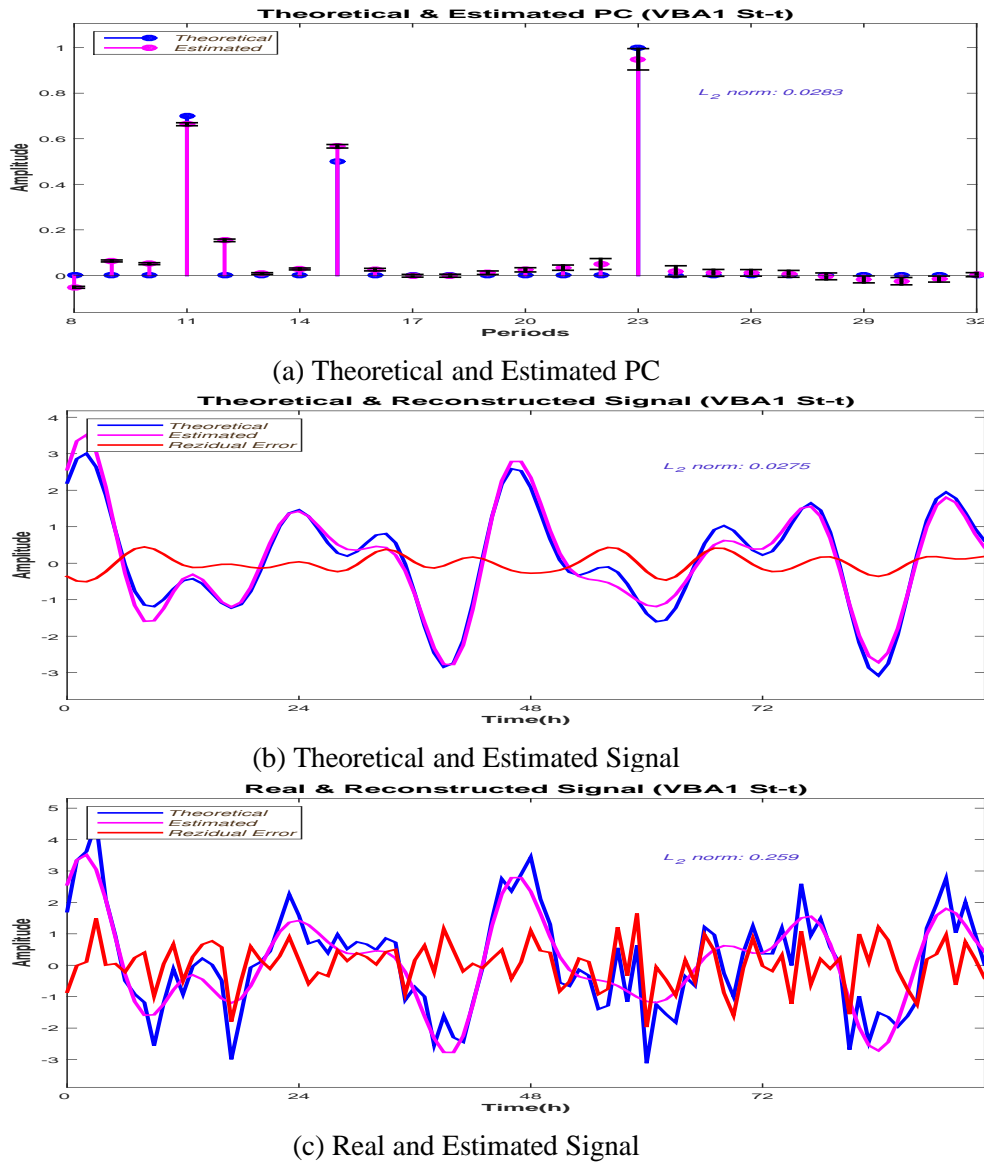


Figure 5.4: PM (via VBA, partial separability) IGSM Estimation (5dB)

both the PC estimation and theoretical signal  $\mathbf{g}_0$  reconstruction are very accurate, Figure (5.4a) and Figure (5.4b). For the reconstruction of the theoretical signal  $\mathbf{g}_0$ , the  $L_2$  error norm is  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{PM}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.0275$ . For the PC vector, the reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{PM}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.0283$ . 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, 5dB, Figure (5.4c). During the

algorithm both Inverse Gammas shape parameters are constant, Equation (4.59c) and Equation (4.59d). We present the convergence of the scale parameters  $\beta_\epsilon$  and  $\beta_f$ , Figure (5.5d) and

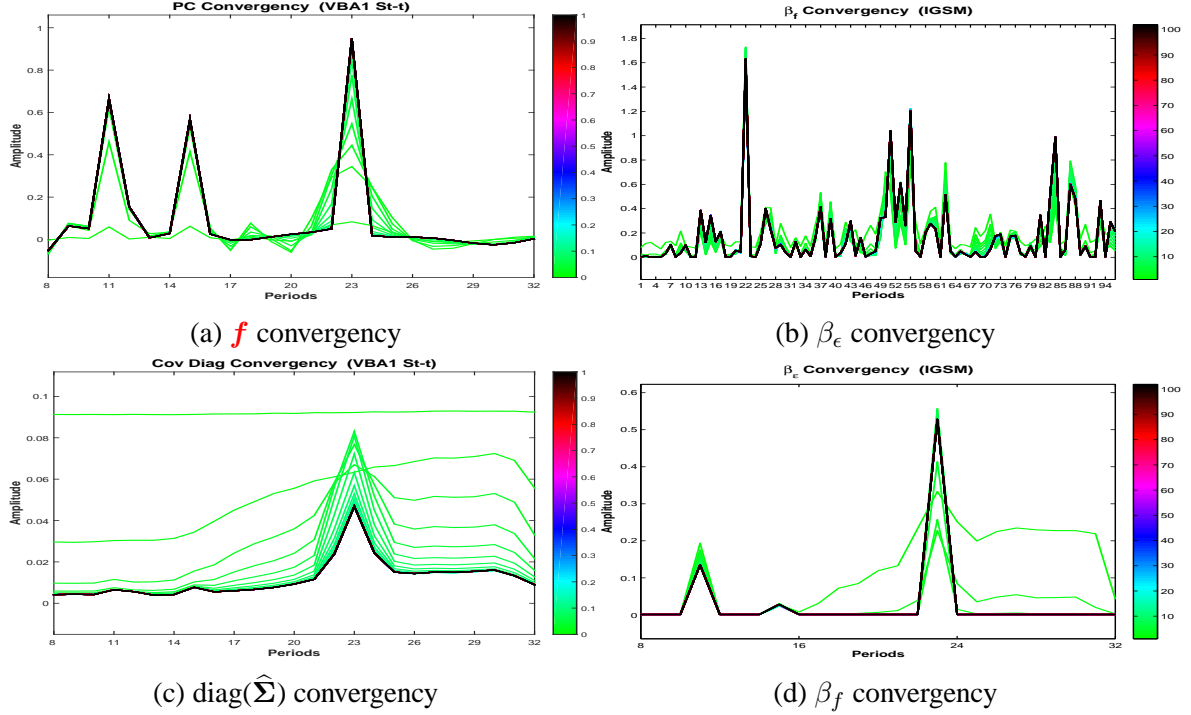


Figure 5.5: PM (via VBA, partial separability) IGSM hyperparameters and  $f$  convergence

Figure (5.5b), the convergence of  $\Sigma$  covariance matrix diagonal, Figure (5.5c) and the convergence of algorithm's solution  $f$ . For a better visualisation of the PC convergence  $f$  is plotted as a vector, Figure (5.5a). The colour 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 (5.6a) presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realisation. 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  $g_0$ , and signal  $g$ , are presented in Figure (5.6b) and Figure (5.6c).

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

The estimations for the full separability case are also accurate: Numerically, for the reconstruction of the theoretical signal  $g_0$ , the  $L_2$  error norm is  $\delta g_0 = \frac{\|g_0 - \hat{g}_{PM}\|_2^2}{\|g_0\|_2^2} = 0.0247$ . For the PC vector, the reconstruction error is  $\delta f = \frac{\|f - \hat{f}_{PM}\|_2^2}{\|f\|_2^2} = 0.0234$ . Figure (5.8a) presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realisation. 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  $g_0$ , and signal  $g$ , are presented in Figure (5.8b) and Figure (5.8c).

### 5.1.6 - Methods comparison 05dB

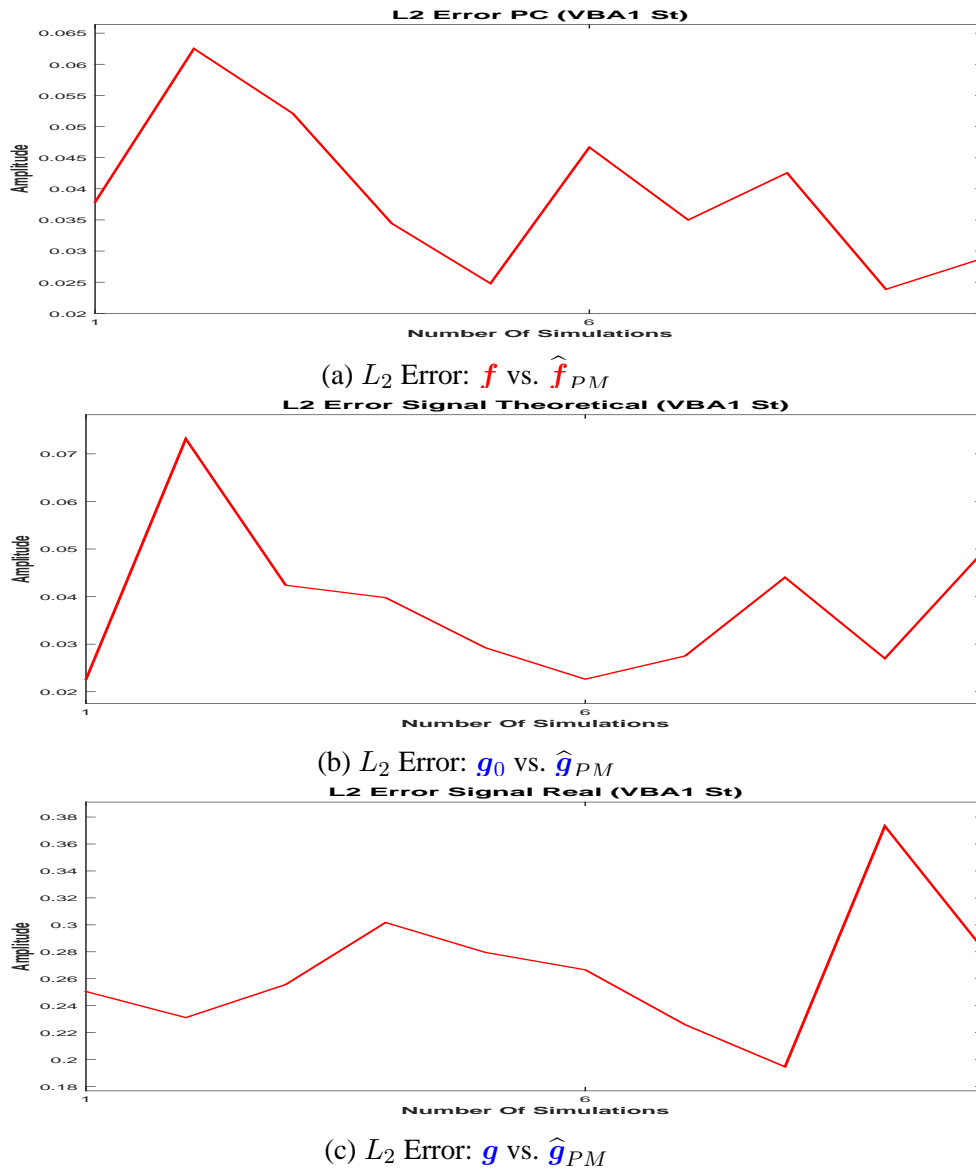
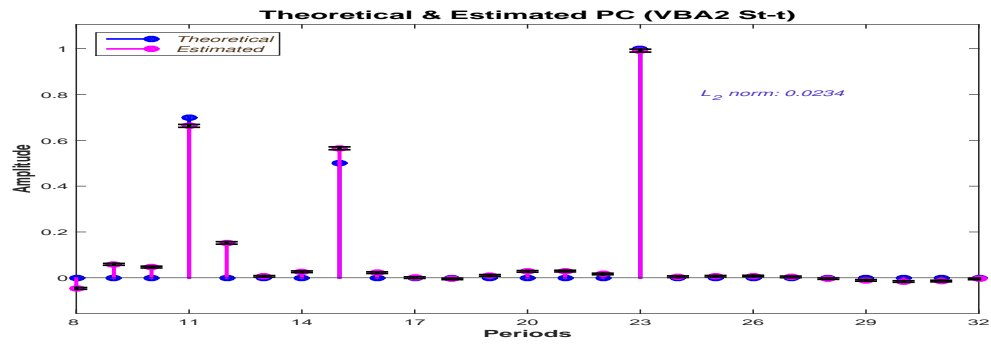


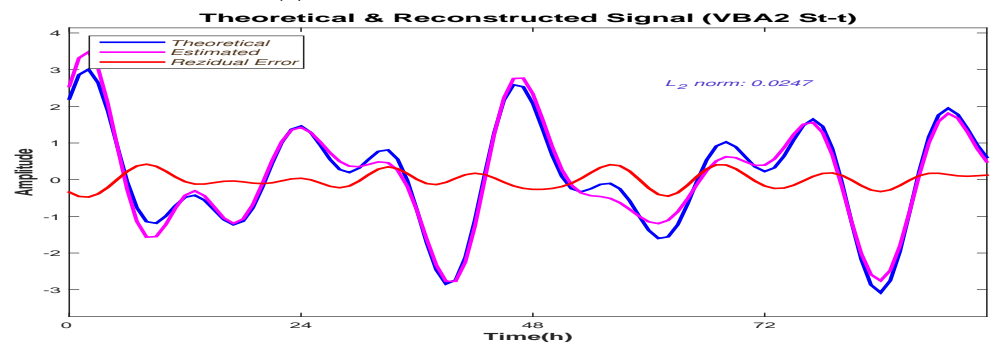
Figure 5.6: PM (via VBA, partial separability) IGSM  $L_2$  error measured for 10 different noise realisations (5dB)

### 5.1.6 Methods comparison 05dB

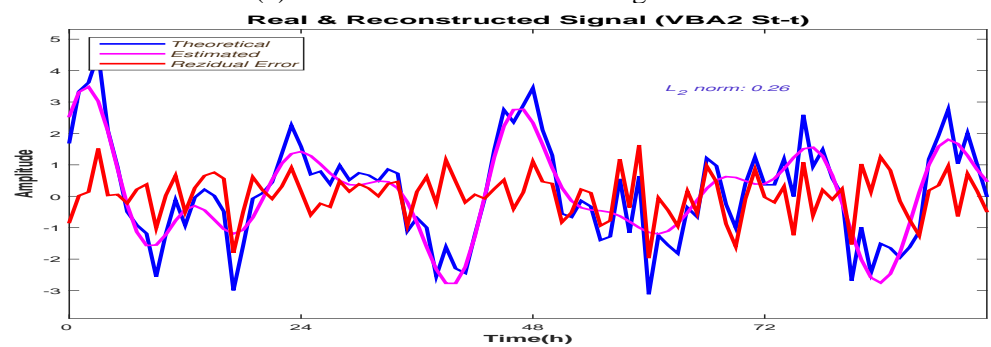
A comparison between the estimations corresponding to the IGSM proposed model is presented in Figure (5.9c) (JMAP estimator), Figure (5.9d) (PM via VBA, partial separability estimator) and Figure (5.9e) (PM via VBA, full separability estimator). As mentioned in Section (4.1), during this article we adopted 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, Figure (5.9a) (Gaussian Model, JMAP estimator) and Figure (5.9b) (PM via VBA estimator). A comparison with the FFT is presented in Figure (5.9f). 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.



(a) Theoretical and Estimated PC



(b) Theoretical and Estimated Signal



(c) Real and Estimated Signal

Figure 5.7: PM (via VBA, full separability) IGSM Estimation (5dB)

### 5.1.7 Error comparison 05dB

The  $L_2$  error measurement corresponding to the PC estimation, theoretical signal estimation and the signal estimation, for 10 different noise realisation is presented in the following figure: The  $L_2$  error corresponding to the PM via VBA IGSM Model corresponding to the PC vector estimation, Figure (5.10a) shows the performances of the proposed algorithm compared to the Gaussian Model and the JMAP estimation for IGSM Model.



### 5.1.7 - Error comparison 05dB

---

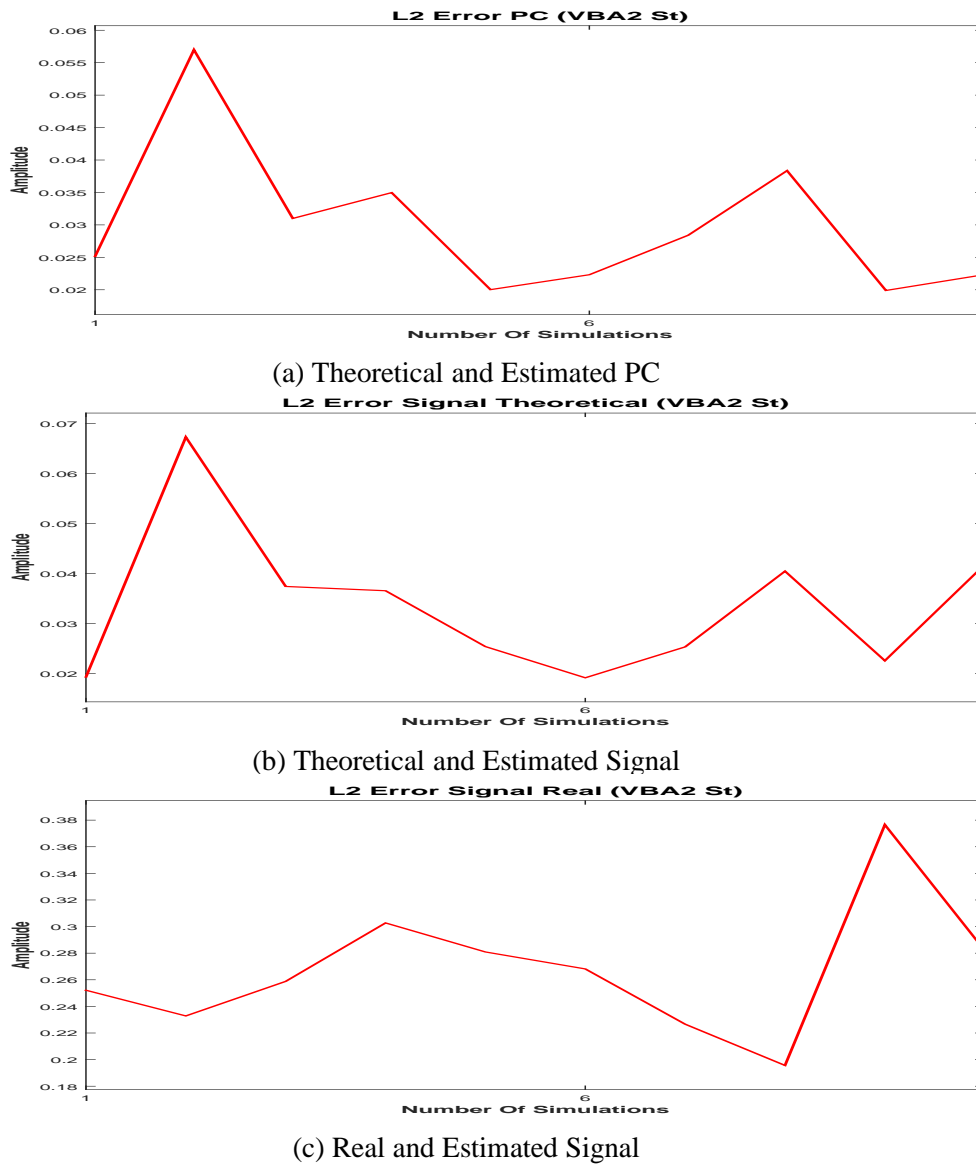


Figure 5.8: PM (via VBA, full separability) IGSM  $L_2$  error measured for 10 different noise realisations (5dB)

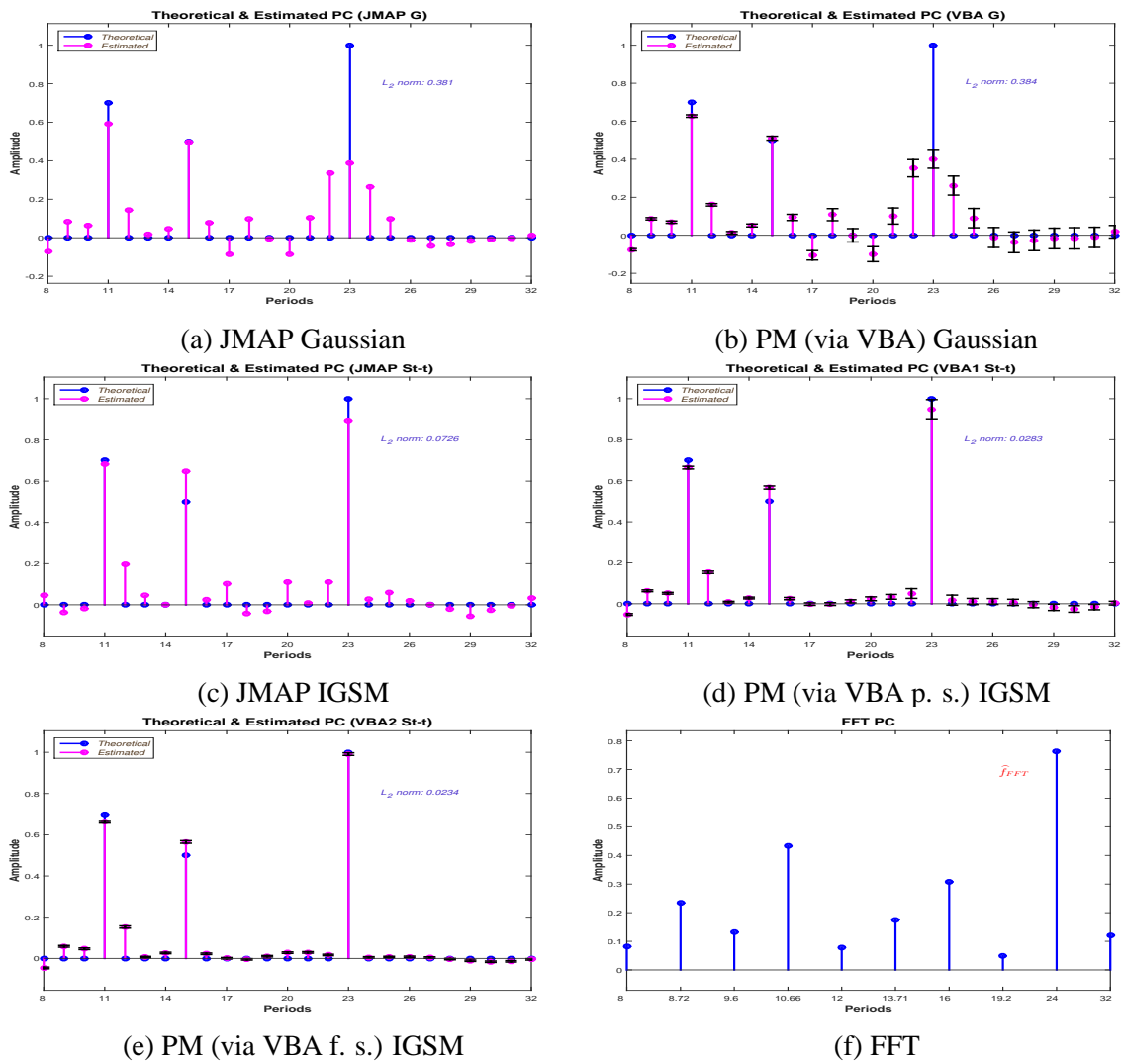
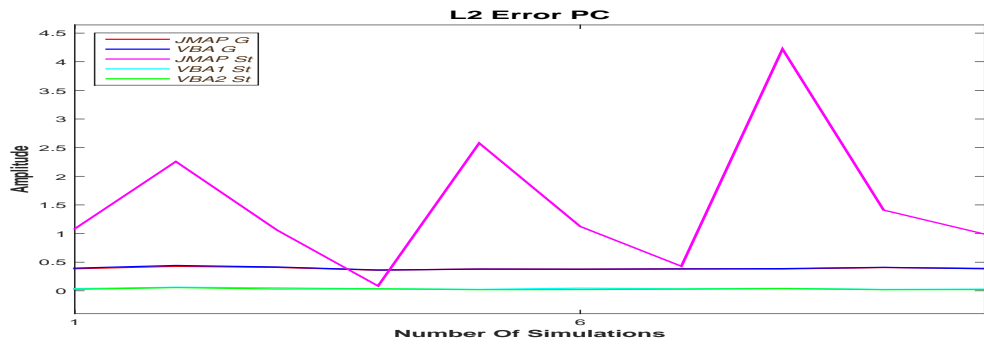
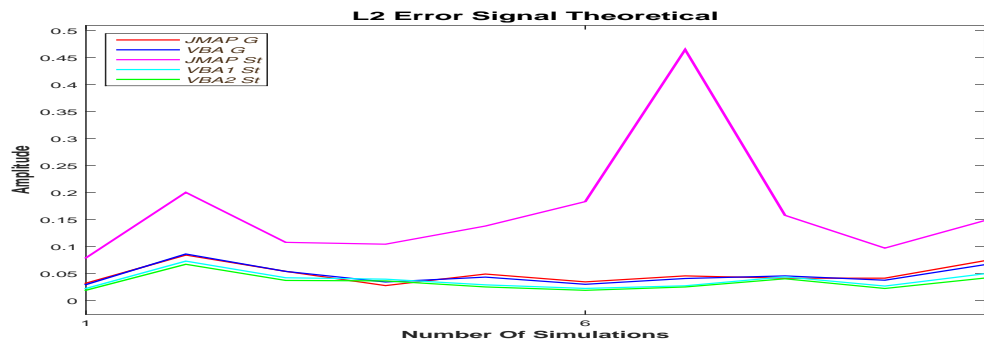


Figure 5.9: Methods Comparison (5dB)

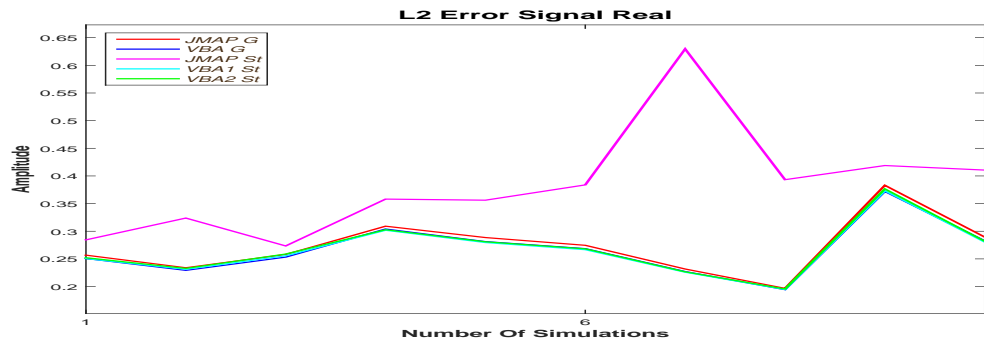
### 5.1.7 - Error comparison 05dB



(a) PC error estimation



(b) Theoretical signal error estimation



(c) Signal error estimation (5dB)

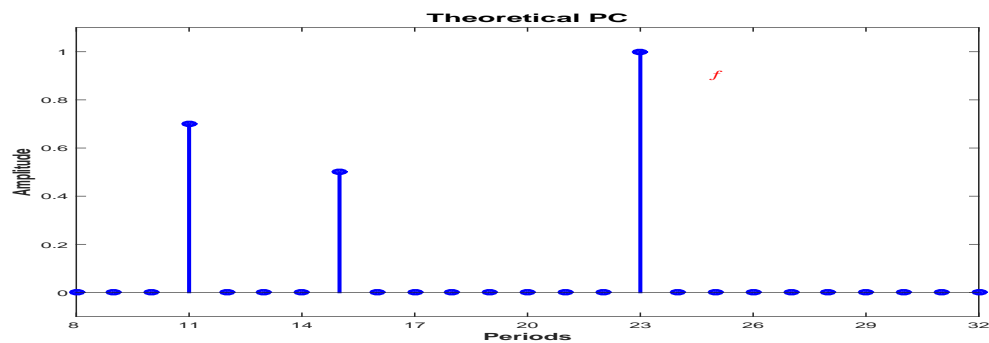
Figure 5.10:  $L_2$  Errors estimation (5dB)

## 5.2 Synthetic data 10dB

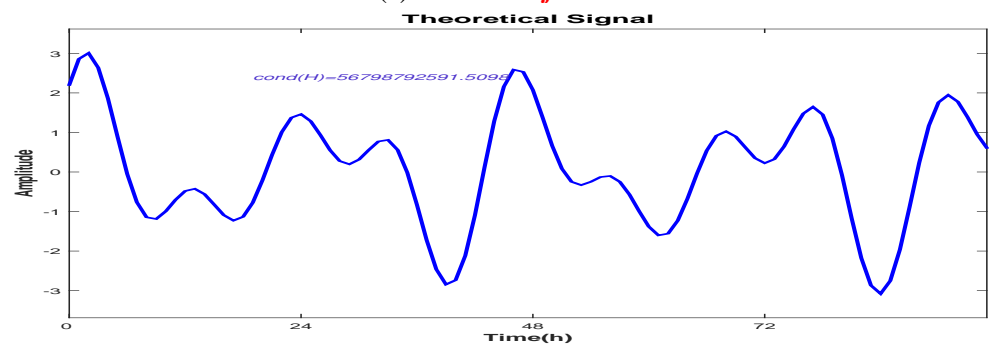
In this subsection we present the synthetic data case corresponding to the 10 dB SNR.

### 5.2.1 Data 10dB

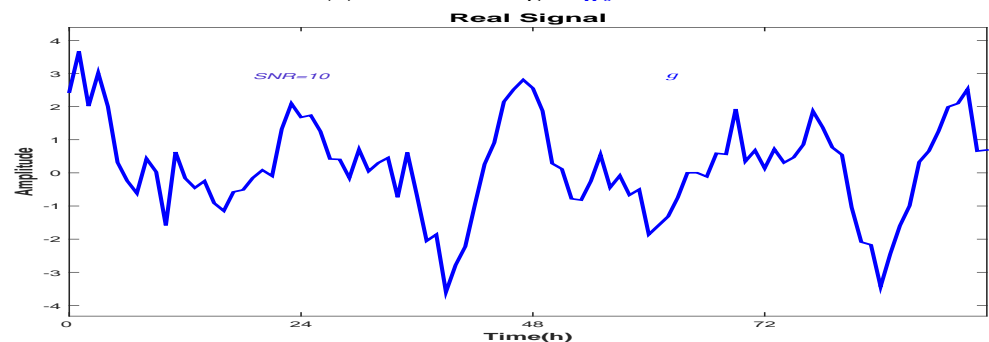
The periodic component vector  $f$ , theoretical signal  $g_0$  and the signal  $g$  are presented in the following figure: Figure (5.11a) shows the theoretical PC, set for 11h, 15h and 23h. All



(a) PC vector  $f$



(b) Theoretical Signal  $g_0$



(c) Signal  $g$

Figure 5.11:  $f$  PC vector, theoretical signal  $g_0$  and input signal  $g = g_0 + \epsilon$  of the model (10dB)

the other values in the PC vector are zero. Figure (5.11b) presents the signal corresponding to the linear model considered in Equation (3), neglecting the errors,  $g_0 = Hf$ . All the simulations are done using as the input the noisy signal  $g$ , (5.11c) corresponding to the linear model, Equation (3).

### 5.2.2 JMAP IGSM 10dB

A comparison between the synthetic data and the JMAP IGSM estimation is presented in the Figure (5.12). The theoretical periodic component vector  $\mathbf{f}$  and the JMAP IGSM estimation  $\hat{\mathbf{f}}_{JMAP}$  are compared in Figure (5.12a). The reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{JMAP}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.0259$ . The comparison between the estimated  $\hat{\mathbf{g}}_{JMAP}$  and the theoretical signal (without noise)  $\mathbf{g}_0$  is presented in in Figure (5.12b). The reconstruction error is  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{JMAP}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.0225$ . The comparison between the estimated  $\hat{\mathbf{g}}_{JMAP}$  and the signal  $\mathbf{g}$  is presented in in Figure (5.12c). Figure (5.13a) presents the variation of  $L_2$  PC vector error reconstruction for

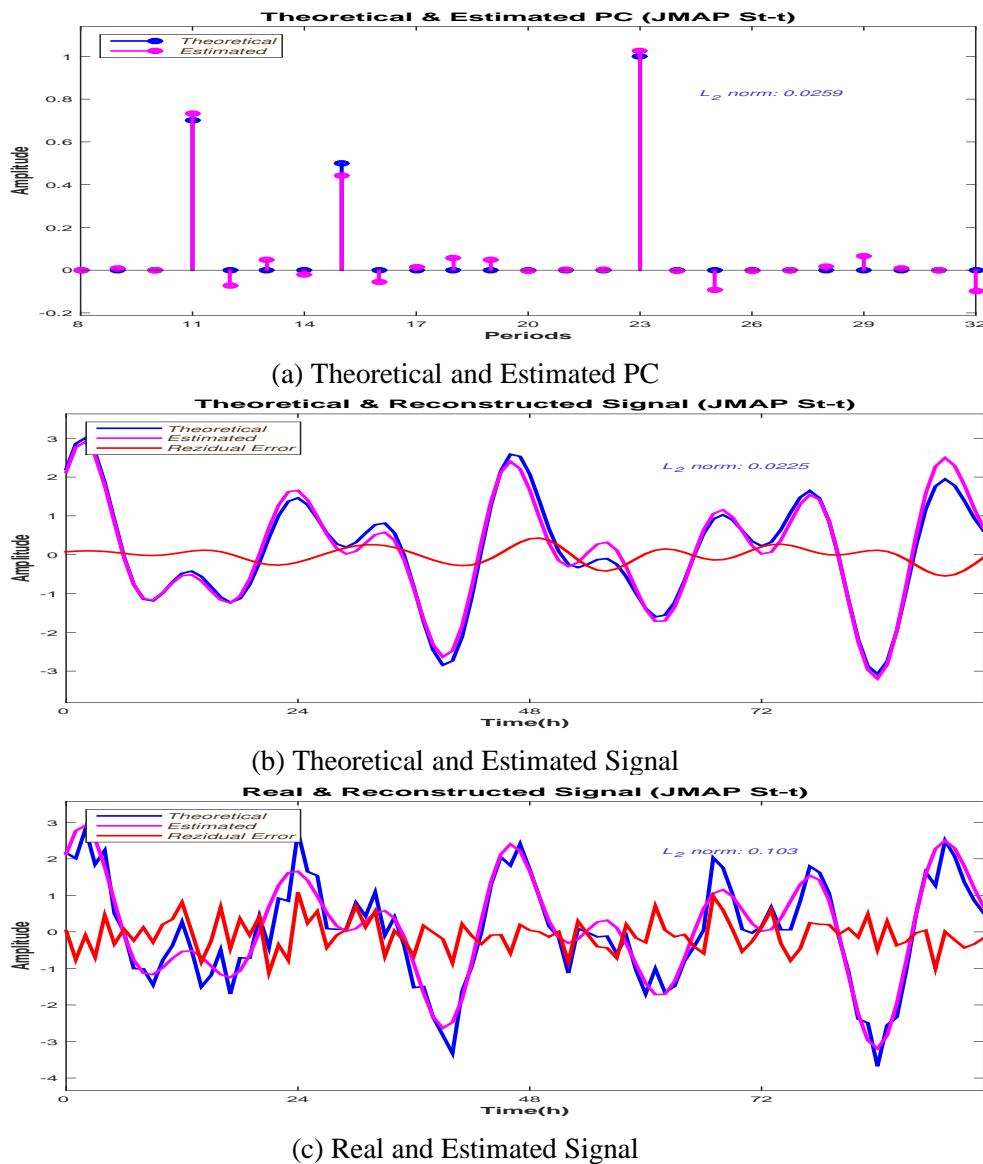


Figure 5.12: JMAP IGSM Estimation (10dB)

10 different noise realisation, showing high variations. Important variations corresponding to the  $L_2$  error reconstruction for the theoretical signal  $\mathbf{g}_0$ , and signal  $\mathbf{g}$ , are presented in Figure (5.13b) and Figure (5.13c).

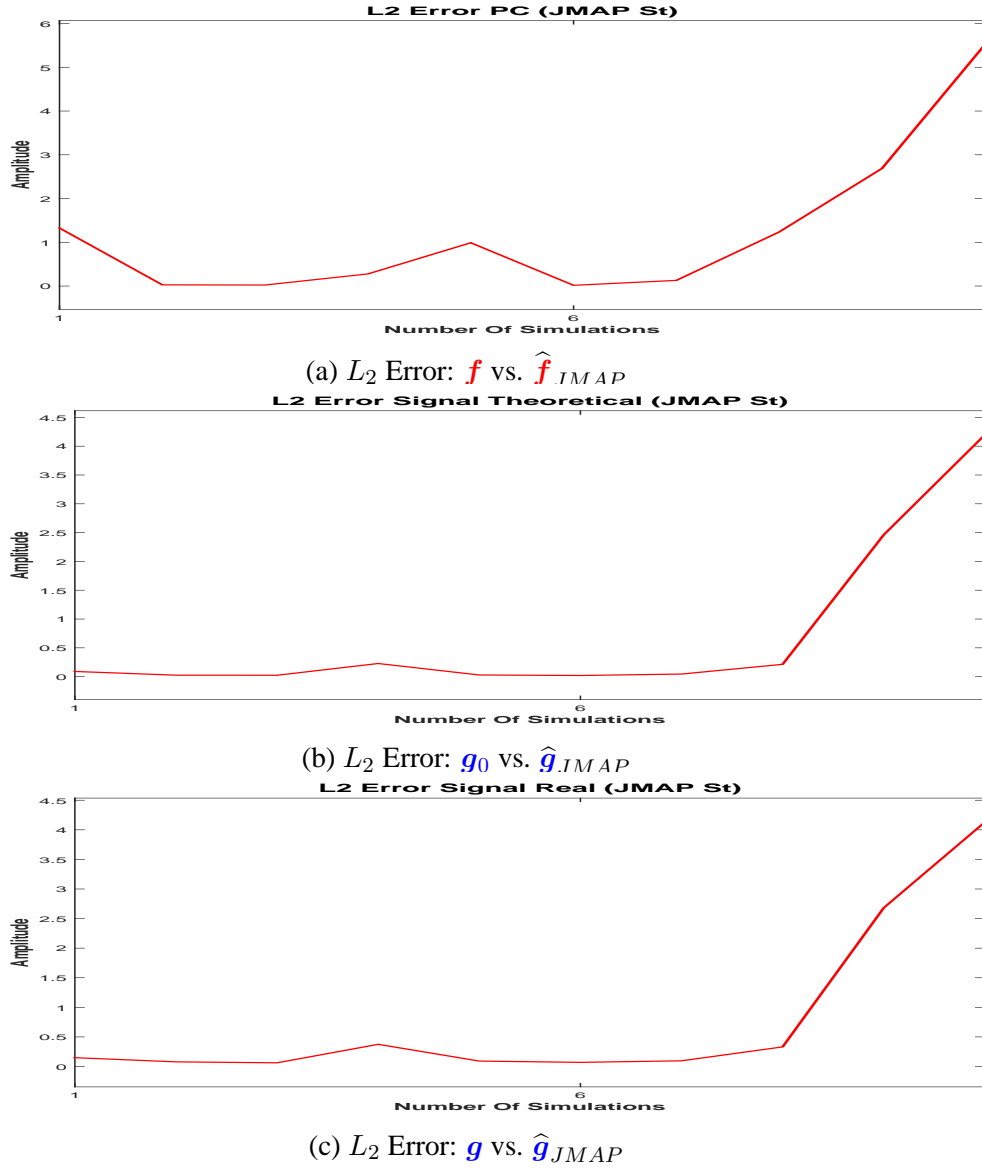


Figure 5.13: JMAP IGSM  $L_2$  error measured for 10 different noise realisations (10dB)

### 5.2.3 PM (via VBA, partial separability) IGSM 10dB

For the Posterior Mean estimation via VBA, both the PC estimation and theoretical signal  $\mathbf{g}_0$  reconstruction are very accurate, Figure (5.14a) and Figure (5.14b). For the reconstruction of the theoretical signal  $\mathbf{g}_0$ , the  $L_2$  error norm is  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{PM}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.011$ . For the PC vector, the reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{PM}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.0144$ . The algorithm is converging to a sparse solution where all the non zero clocks are detected. The residual error computed between  $\mathbf{g}$  and the reconstructed signal is consistent with the error considered in the model, 10dB, Figure (5.14c). The convergency of  $\Sigma$  covariance matrix diagonal, Figure (5.15a) and the convergency of algorithm's solution  $\mathbf{f}$ , Figure (5.15b) is showing a very fast convergency. All the estimates are superposed after the first ten iterations. We show that for the PM estimation, the error variation is very small. Figure (5.16a) presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realisation. The figure presents a very small variation of  $L_2$  PC vector error reconstruction, between 0.008 and 0.02. Very small variations

### 5.2.3 - PM (via VBA, partial separability) IGSM 10dB

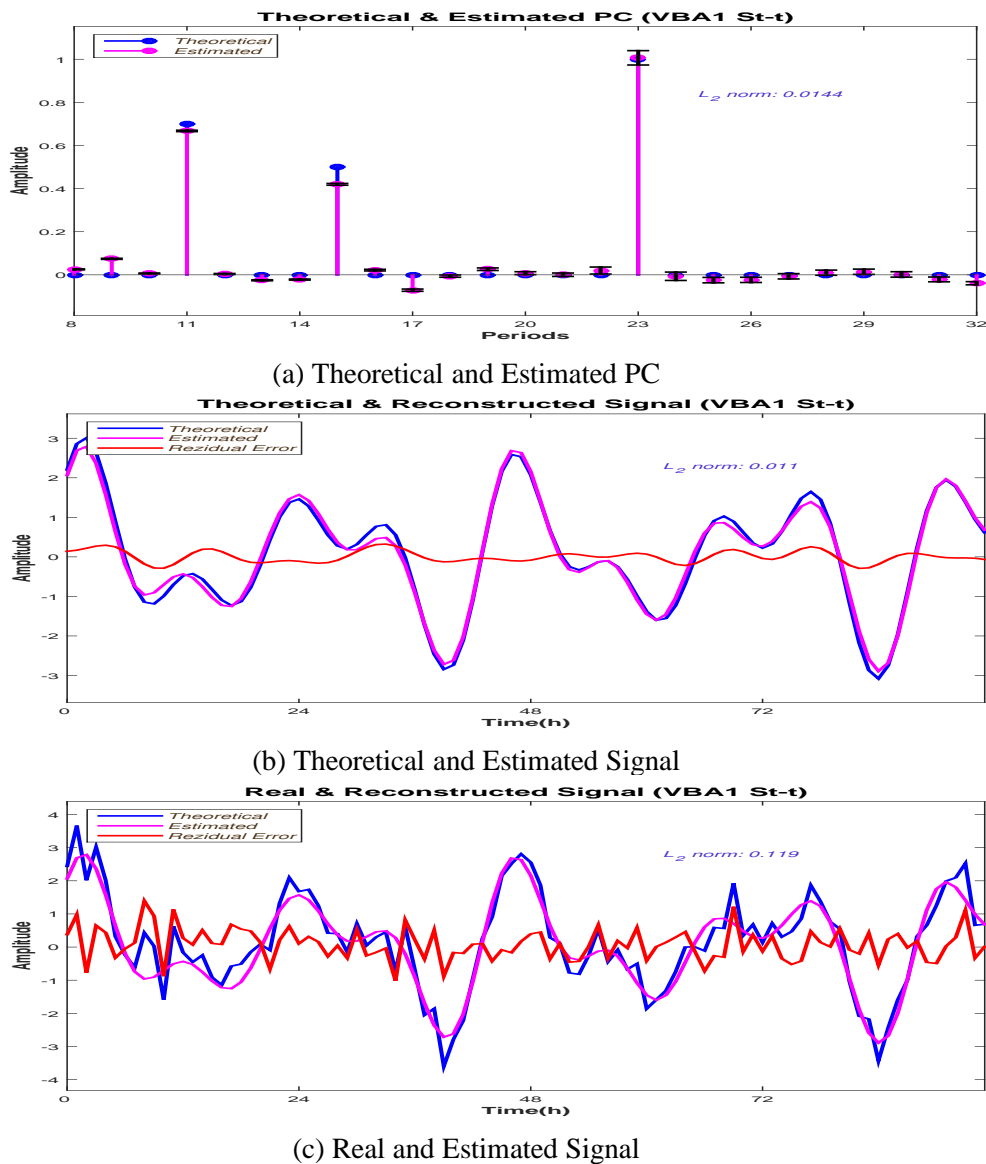


Figure 5.14: PM (via VBA, partial separability) IGSM Estimation (10dB)

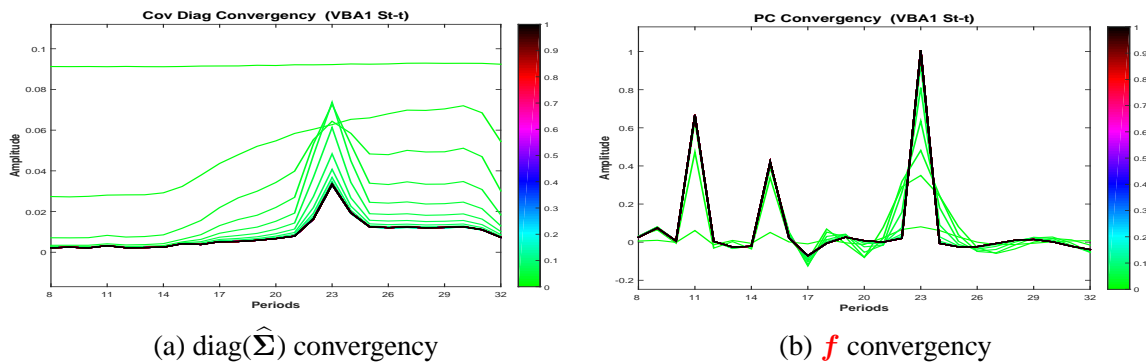


Figure 5.15: PM (via VBA, partial separability) IGSM hyperparameters and  $f$  convergency

corresponding to the  $L_2$  error reconstruction for the theoretical signal  $g_0$ , and signal  $g$ , are presented in Figure (5.16b) and Figure (5.16c).

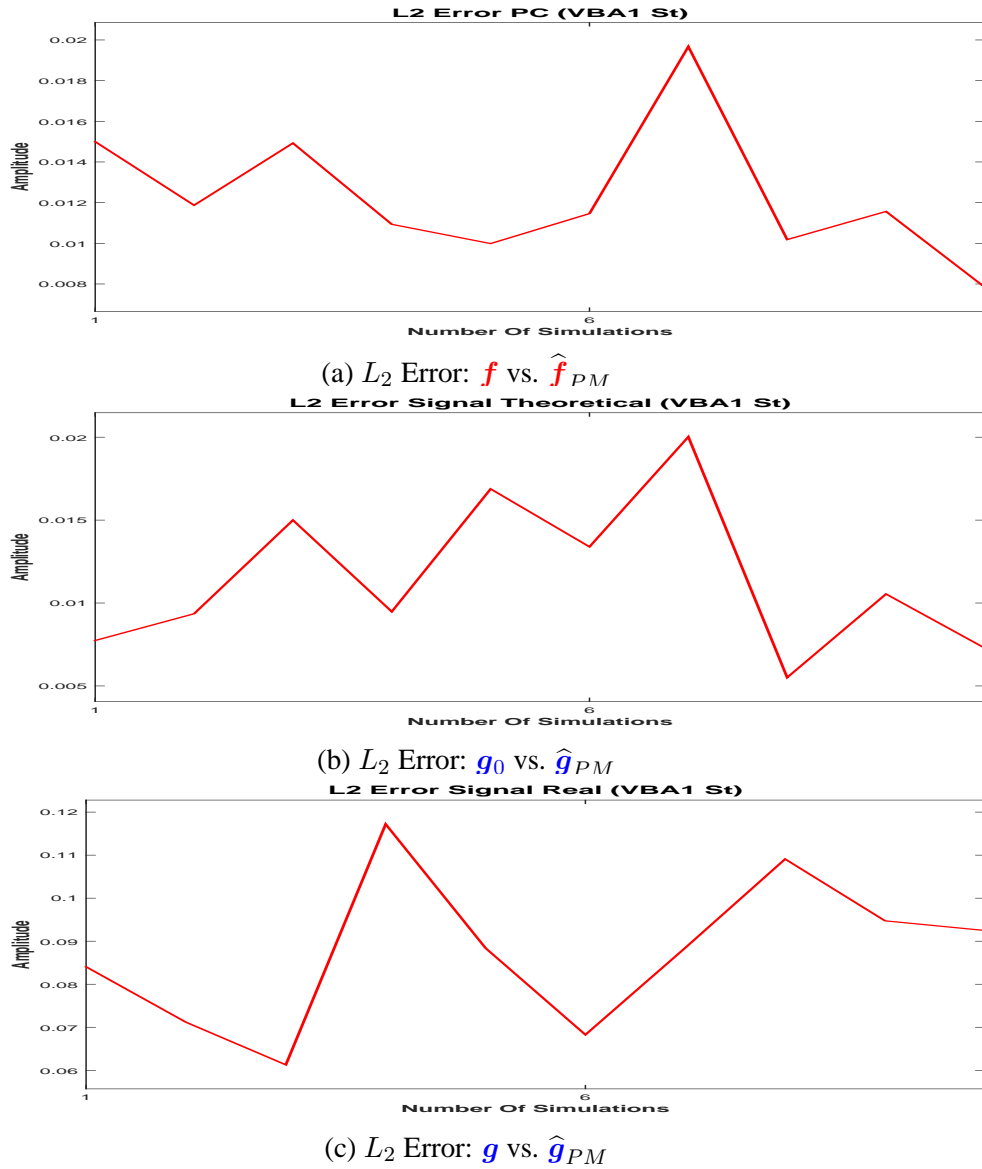


Figure 5.16: PM (via VBA, partial separability) IGSM  $L_2$  error measured for 10 different noise realisations (10dB)

## 5.2.4 PM (via VBA, full separability) IGSM 10dB

The estimations for the full separability case are also accurate: For the reconstruction of the theoretical signal  $\mathbf{g}_0$ , the  $L_2$  error norm is  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{PM}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.0095$ . For the PC vector, the reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{PM}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.0103$ . Figure (5.18a) presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realisation. The figure presents a very small variation of  $L_2$  PC vector error reconstruction, between 0.0046 and 0.0168. Very small variations corresponding to the  $L_2$  error reconstruction for the theoretical signal  $\mathbf{g}_0$ , and signal  $\mathbf{g}$ , are presented in Figure (5.18b) and Figure (5.18c).



## 5.2.5 - Methods comparison 10dB

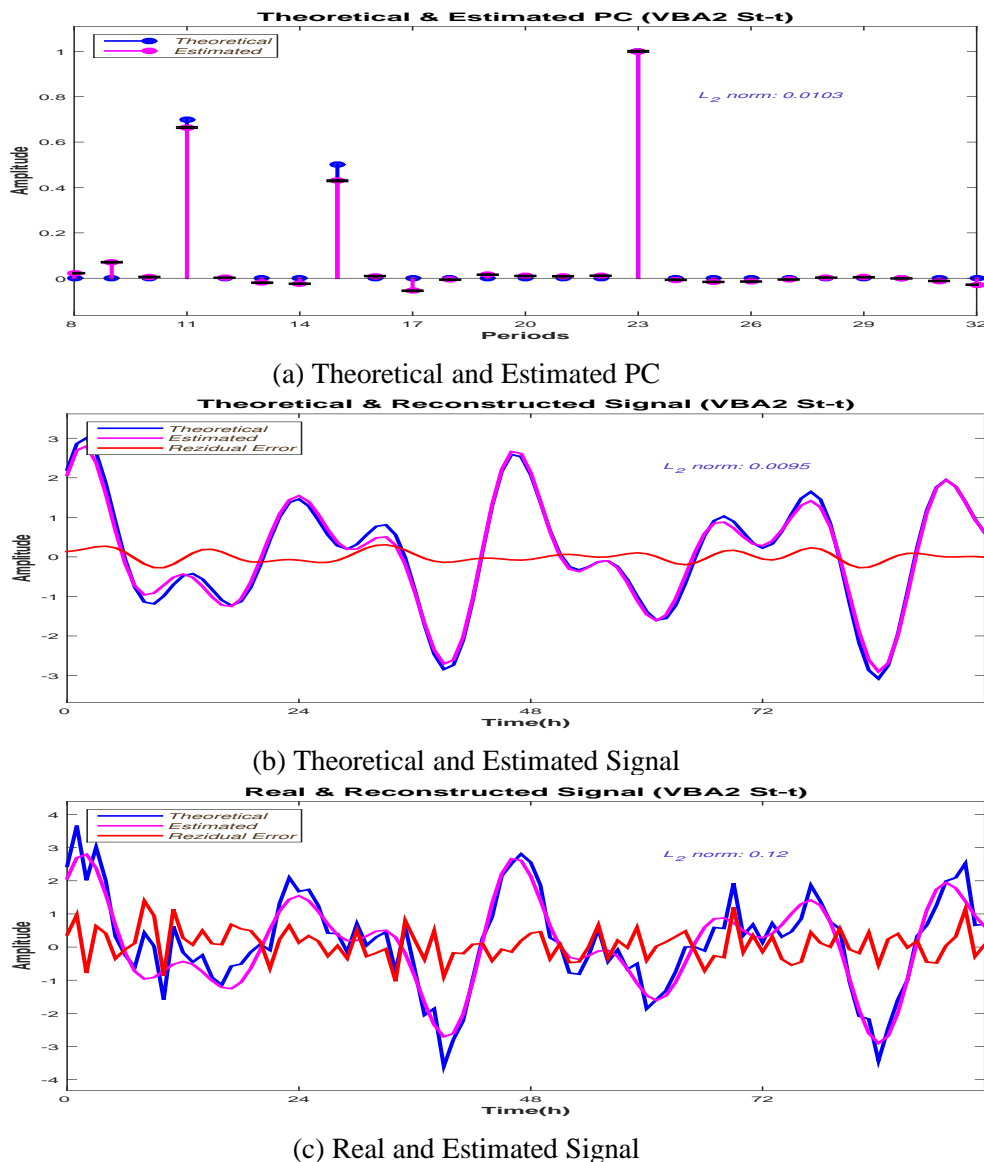
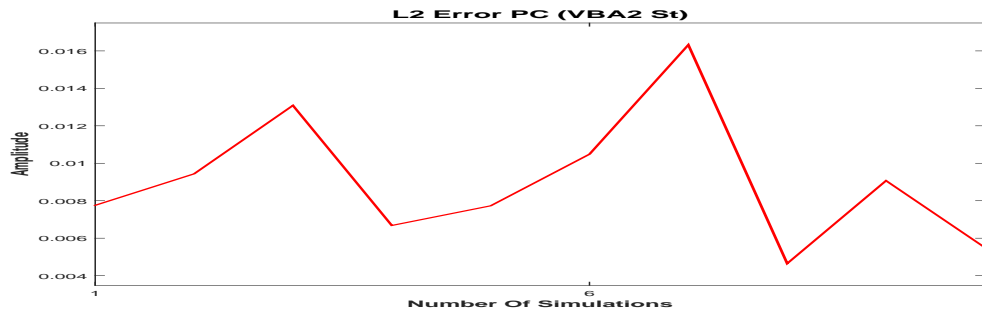


Figure 5.17: PM (via VBA, full separability) IGSM Estimation (10dB)

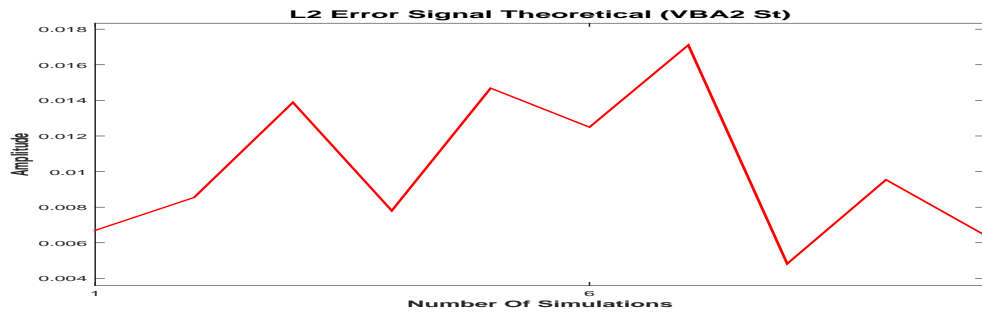
## 5.2.5 Methods comparison 10dB

The comparison between the estimations corresponding to the proposed IGSM models is presented in Figure (5.19c) (JMAP estimator), Figure (5.19d) (PM via VBA, partial separability estimator) and Figure (5.19e) (PM via VBA, full separability estimator). The comparison with the Gaussian case (i.e. Gaussian prior), via the two estimators discussed, is presented in Figure (5.19a) (Gaussian Model, JMAP estimator) and Figure (5.19b) (PM via VBA estimator). A comparison with the FFT is presented in Figure (5.19f).

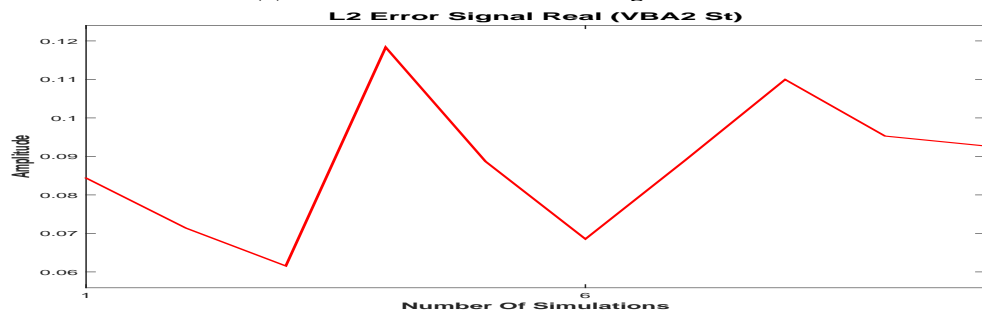
The  $L_2$  estimation error for the PC vector is very high for the two Gaussian models. Also, the estimations are not sparse. For the IGSM models, the JMAP estimator is providing a good estimation, but it is unstable. VBA estimator, both partial separable and fully separable provides very accurate stable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



(a) Theoretical and Estimated PC



(b) Theoretical and Estimated Signal



(c) Real and Estimated Signal

Figure 5.18: PM (via VBA, full separability) IGSM  $L_2$  error measured for 10 different noise realisations (10dB)

### 5.2.6 Error comparison 10dB

The  $L_2$  error measurement corresponding to the PC estimation, theoretical signal estimation and the signal estimation, for 10 different noise realisation is presented in the following figure: The  $L_2$  error corresponding to the PM via VBA IGSM Model shows the performances of proposed via algorithm compared to the Gaussian Model and the JMAP estimation for IGSM Model.

## 5.2.6 - Error comparison 10dB

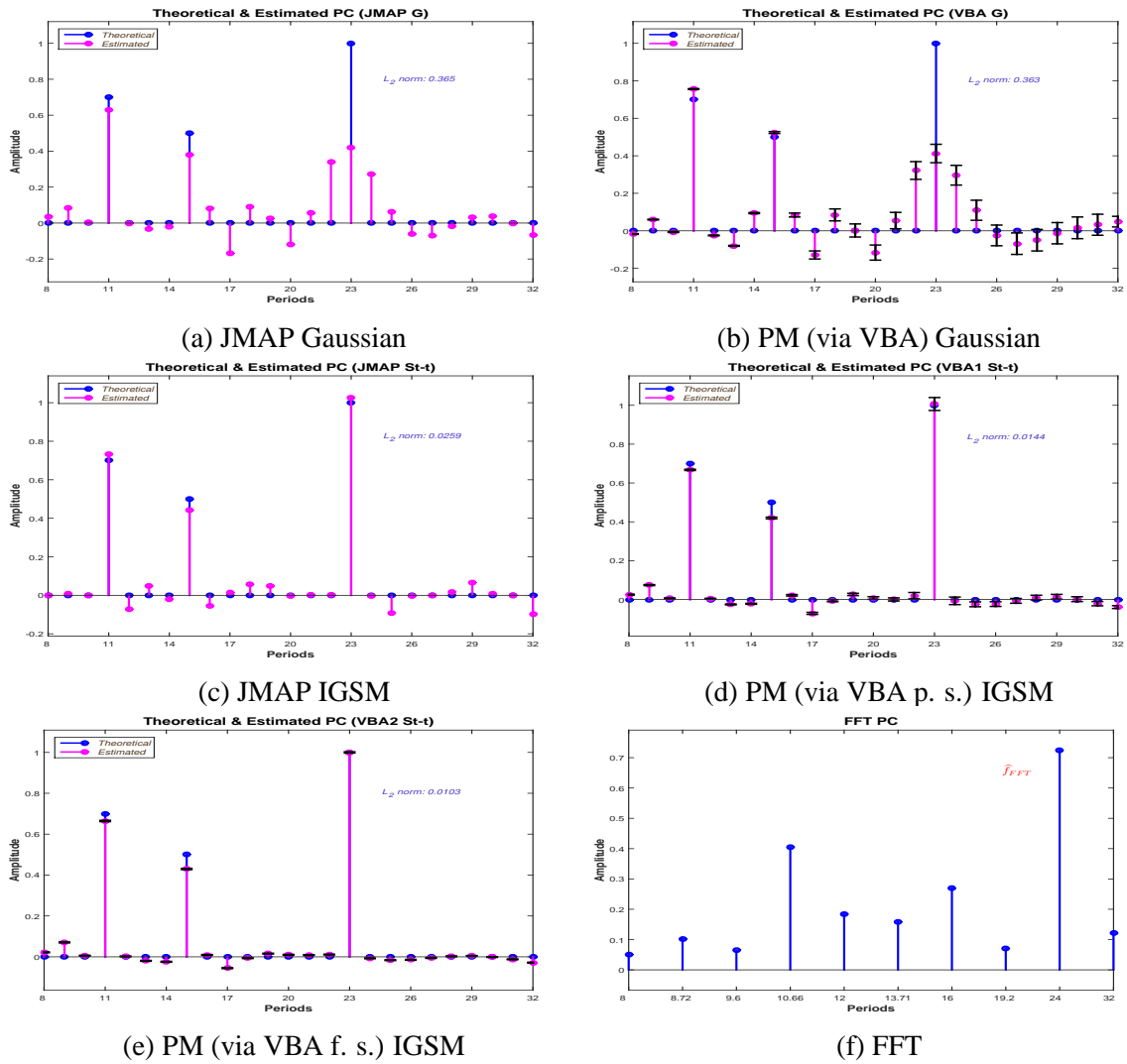
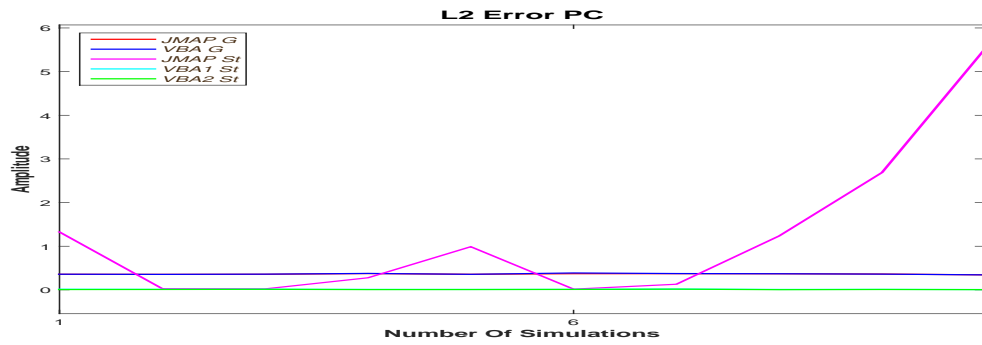
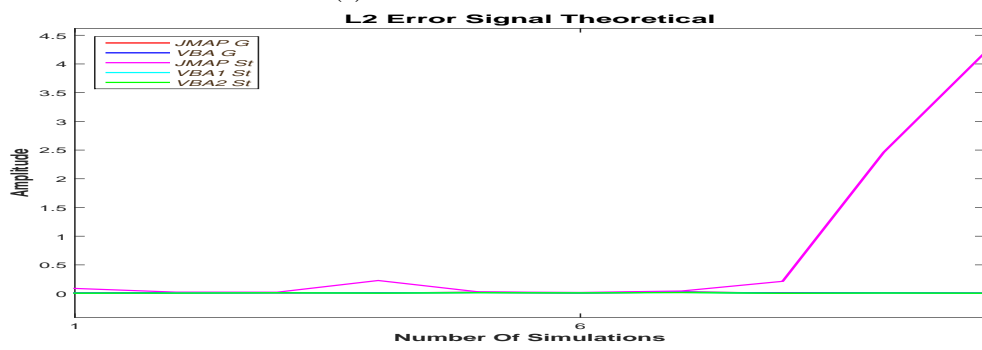


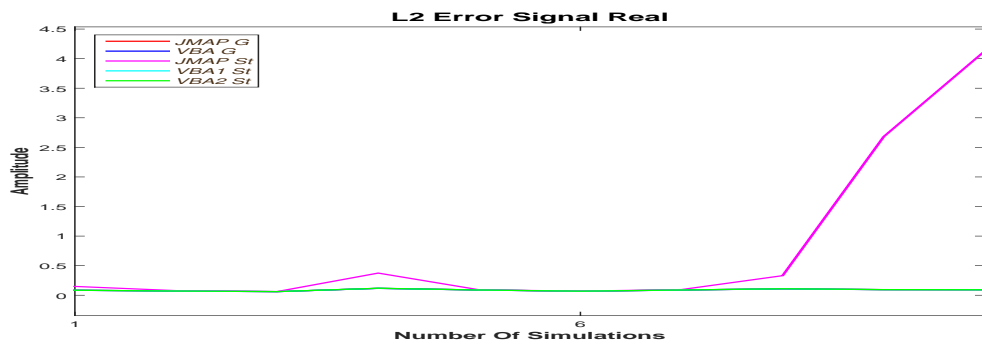
Figure 5.19: Methods Comparison (10dB)



(a) PC error estimation



(b) Theoretical signal error estimation



(c) Signal error estimation (10dB)

Figure 5.20:  $L_2$  Errors estimation (10dB)

## 5.3 Synthetic data 15dB

In this subsection we present the synthetic data case corresponding to the 15 dB SNR.

### 5.3.1 Data 15dB

The periodic component vector  $f$ , theoretical signal  $g_0$  and the signal  $g$  are presented in the following figure: Figure (5.21a) shows the theoretical PC, set for 11h, 15h and 23h. All

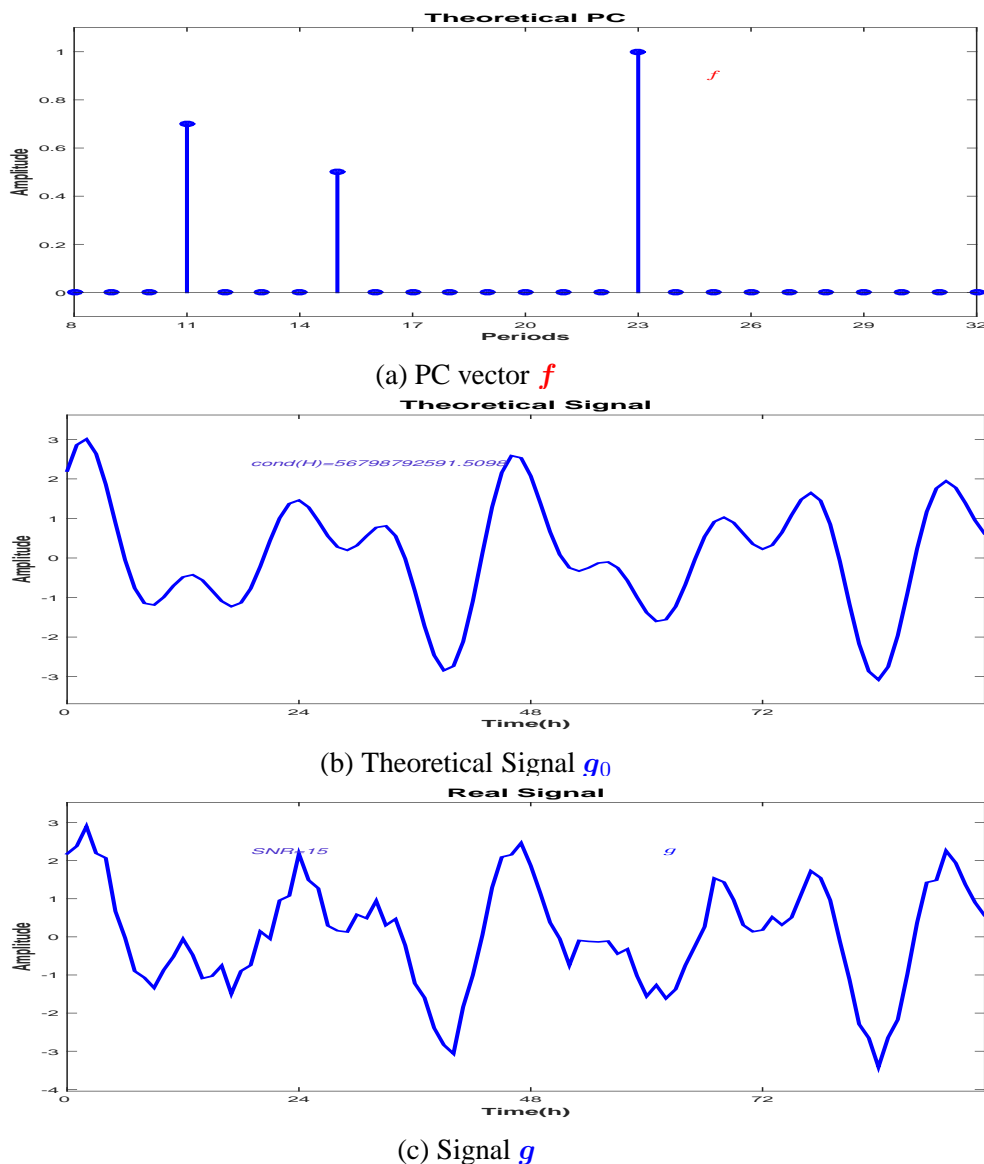


Figure 5.21:  $f$  PC vector, theoretical signal  $g_0$  and input signal  $g = g_0 + \epsilon$  of the model (15dB)

the other values in the PC vector are zero. Figure (5.21b) presents the signal corresponding to the linear model considered in Equation (3), neglecting the errors,  $g_0 = Hf$ . All the simulations are done using as the input the noisy signal  $g$ , (5.21c) corresponding to the linear model, Equation (3).

### 5.3.2 JMAP IGSM 15dB

A comparison between the synthetic data and the JMAP IGSM estimation is presented in the Figure (5.22). The theoretical periodic component vector  $\mathbf{f}$  and the JMAP IGSM estimation  $\hat{\mathbf{f}}_{JMAP}$  are compared in Figure (5.22a). The reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{JMAP}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.0069$ . The comparison between the estimated  $\hat{\mathbf{g}}_{JMAP}$  and the theoretical signal (without noise)  $\mathbf{g}_0$  is presented in in Figure (5.22b). The reconstruction error is  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{JMAP}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.00692$ . The comparison between the estimated  $\hat{\mathbf{g}}_{JMAP}$  and the signal  $\mathbf{g}$  is presented in in Figure (5.22c). Figure (5.23a) presents the variation of  $L_2$  PC vector error reconstruction for

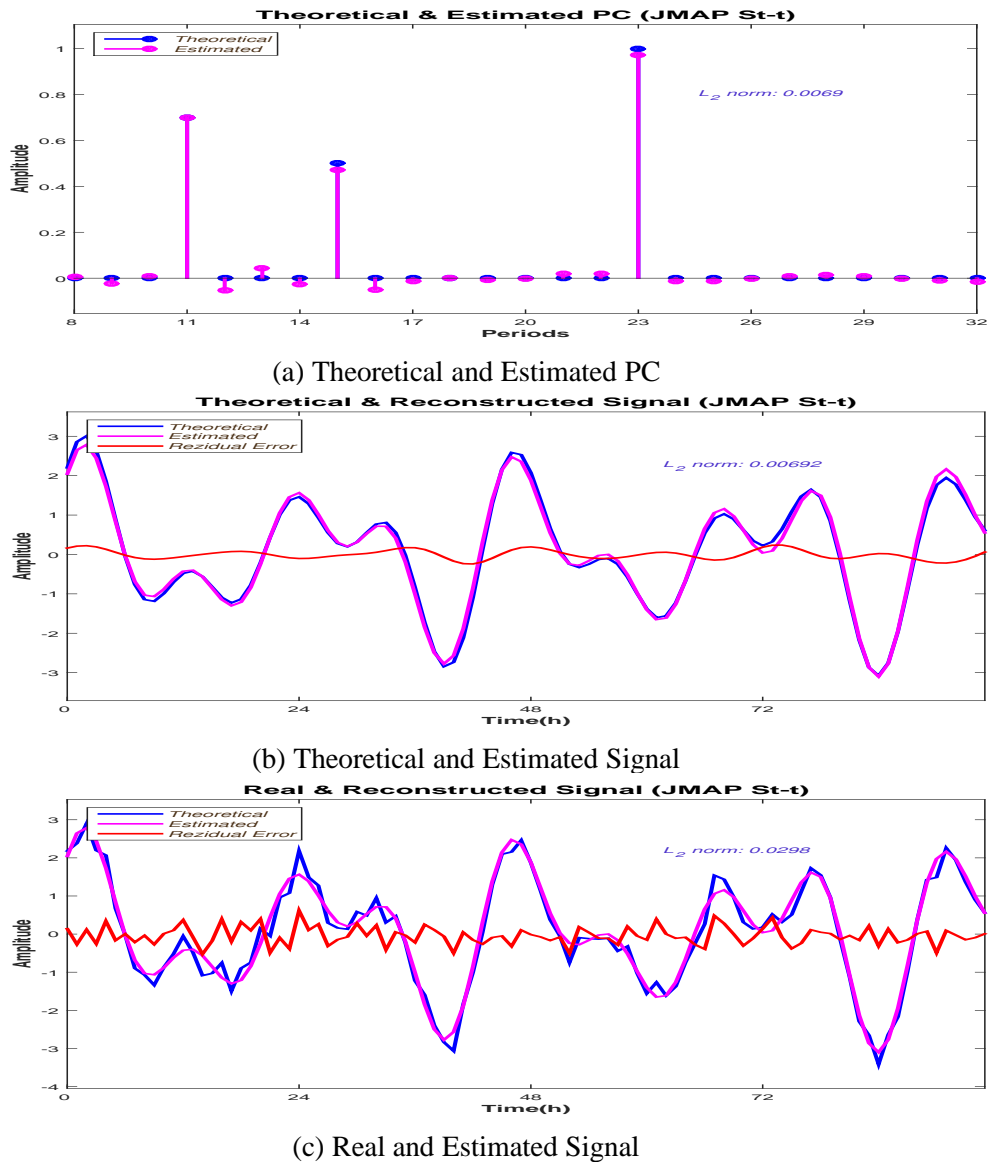


Figure 5.22: JMAP IGSM Estimation (15dB)

10 different noise realisation, showing high variations. Important variations corresponding to the  $L_2$  error reconstruction for the theoretical signal  $\mathbf{g}_0$ , and signal  $\mathbf{g}$ , are presented in Figure (5.23b) and Figure (5.23c).

### 5.3.3 - PM (via VBA, partial separability) IGSM 15dB

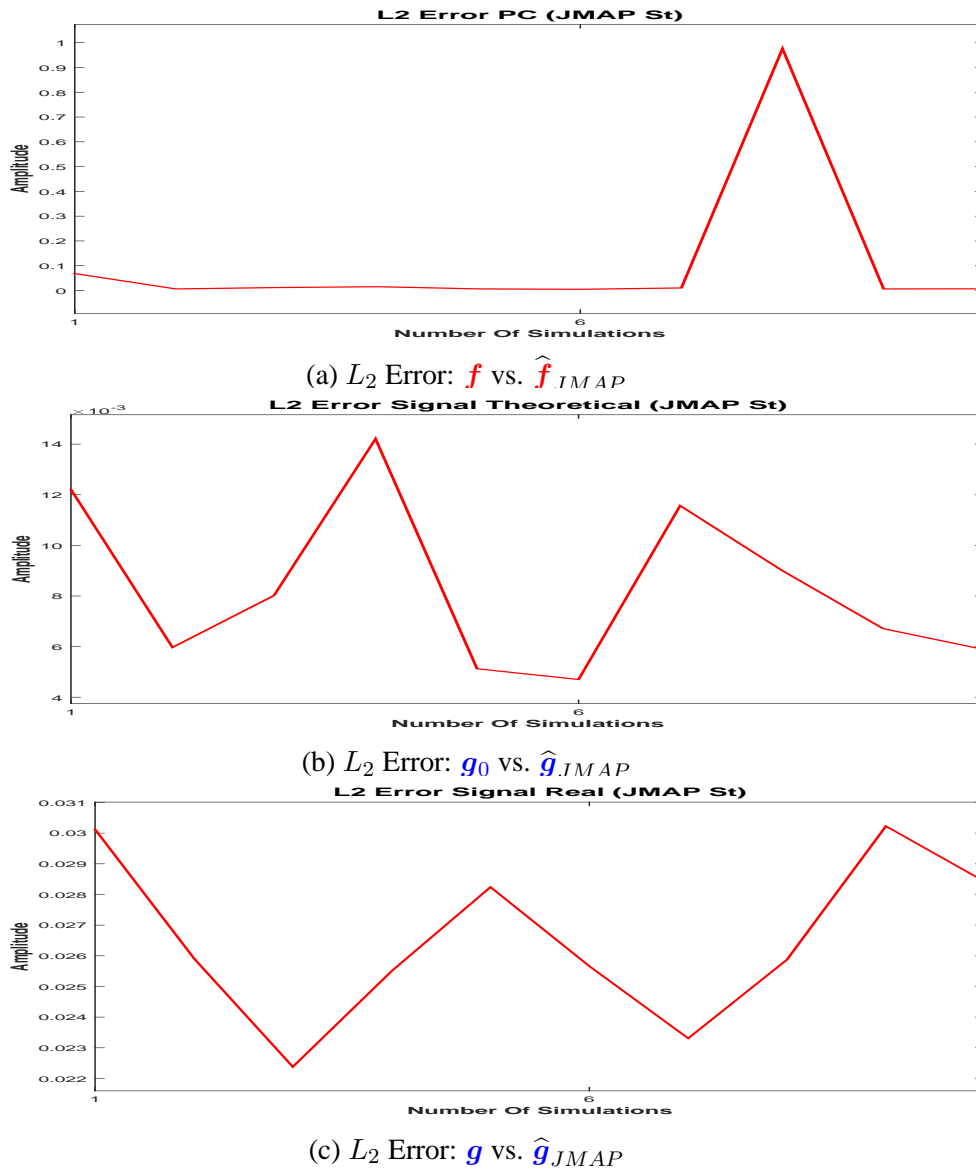
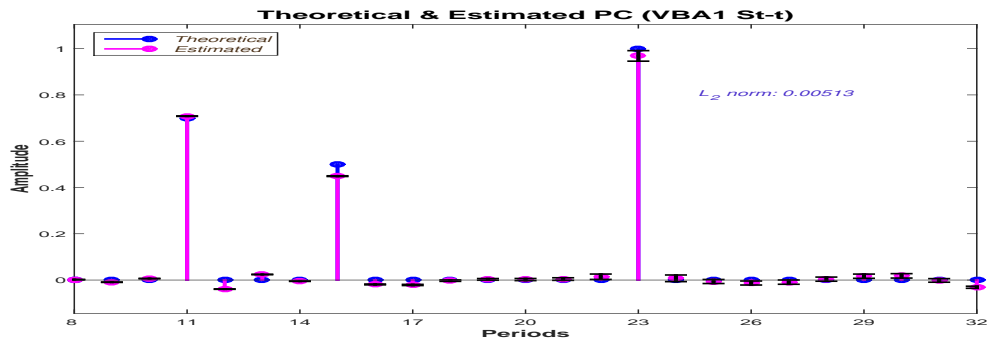


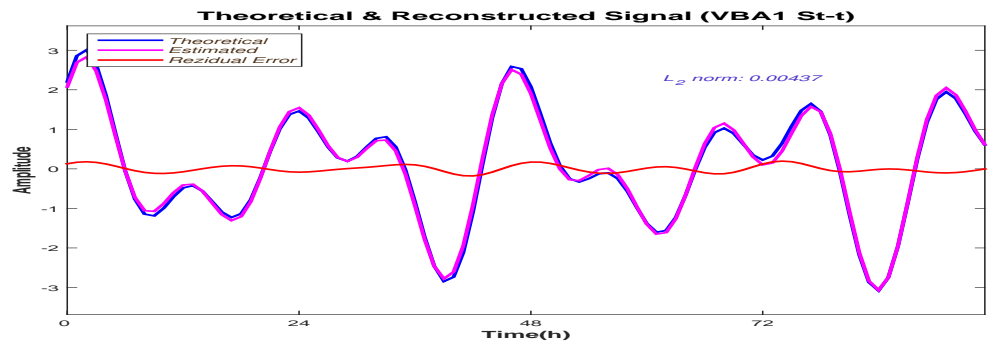
Figure 5.23: JMAP IGSM  $L_2$  error measured for 10 different noise realisations (15dB)

### 5.3.3 PM (via VBA, partial separability) IGSM 15dB

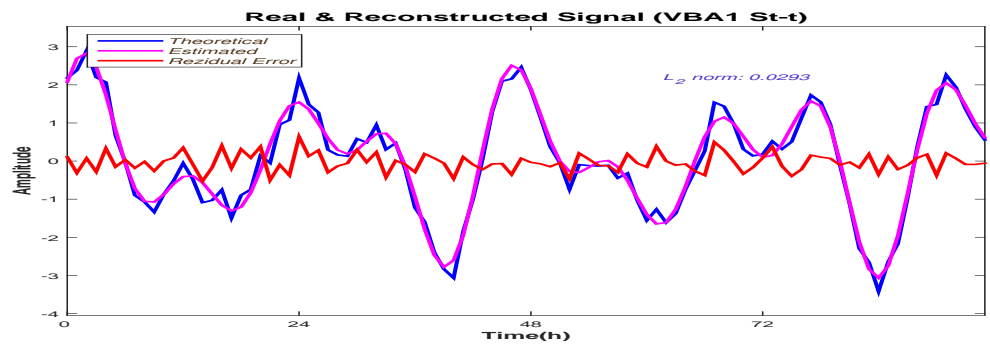
For the Posterior Mean estimation via VBA, both the PC estimation and theoretical signal  $\mathbf{g}_0$  reconstruction are very accurate, Figure (5.24a) and Figure (5.24b). For the reconstruction of the theoretical signal  $\mathbf{g}_0$ , the  $L_2$  error norm is  $\delta \mathbf{g}_0 = \frac{\|\mathbf{g}_0 - \hat{\mathbf{g}}_{PM}\|_2^2}{\|\mathbf{g}_0\|_2^2} = 0.00513$ . For the PC vector, the reconstruction error is  $\delta \mathbf{f} = \frac{\|\mathbf{f} - \hat{\mathbf{f}}_{PM}\|_2^2}{\|\mathbf{f}\|_2^2} = 0.00437$ . The algorithm is converging to a sparse solution where all the non zero clocks are detected. The residual error computed between  $\mathbf{g}$  and the reconstructed signal is consistent with the error considered in the model, 15dB, Figure (5.24c). The convergency of  $\Sigma$  covariance matrix diagonal, Figure (5.25a) and the convergency of algorithm's solution  $\mathbf{f}$ , Figure (5.25b) is showing a very fast convergency. All the estimates are superposed after the first ten iterations. We show that for the PM estimation, the error variation is very small. Figure (5.26a) presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realisation. The figure presents a very small variation of  $L_2$  PC vector error reconstruction, between 0.003 and 0.01. Very small variations



(a) Theoretical and Estimated PC



(b) Theoretical and Estimated Signal



(c) Real and Estimated Signal

Figure 5.24: PM (via VBA, partial separability) IGSM Estimation (15dB)

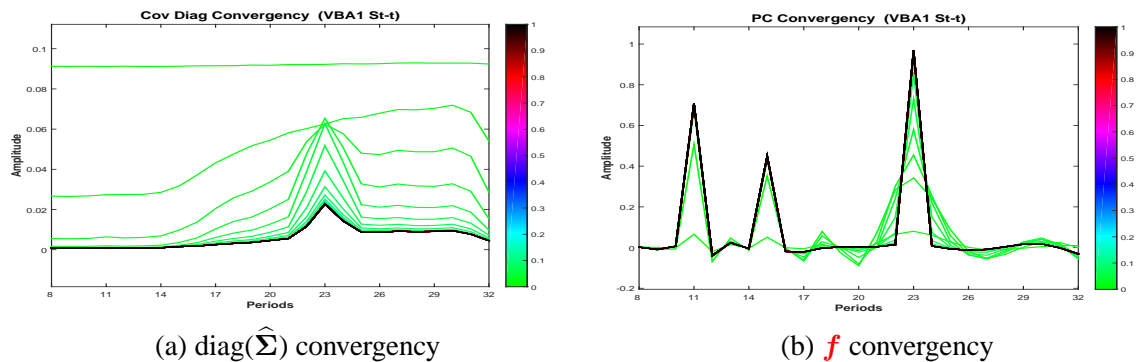


Figure 5.25: PM (via VBA, partial separability) IGSM hyperparameters and  $f$  convergency

corresponding to the  $L_2$  error reconstruction for the theoretical signal  $g_0$ , and signal  $g$ , are presented in Figure (5.26b) and Figure (5.26c).



### 5.3.4 - PM (via VBA, full separability) IGSM 15dB

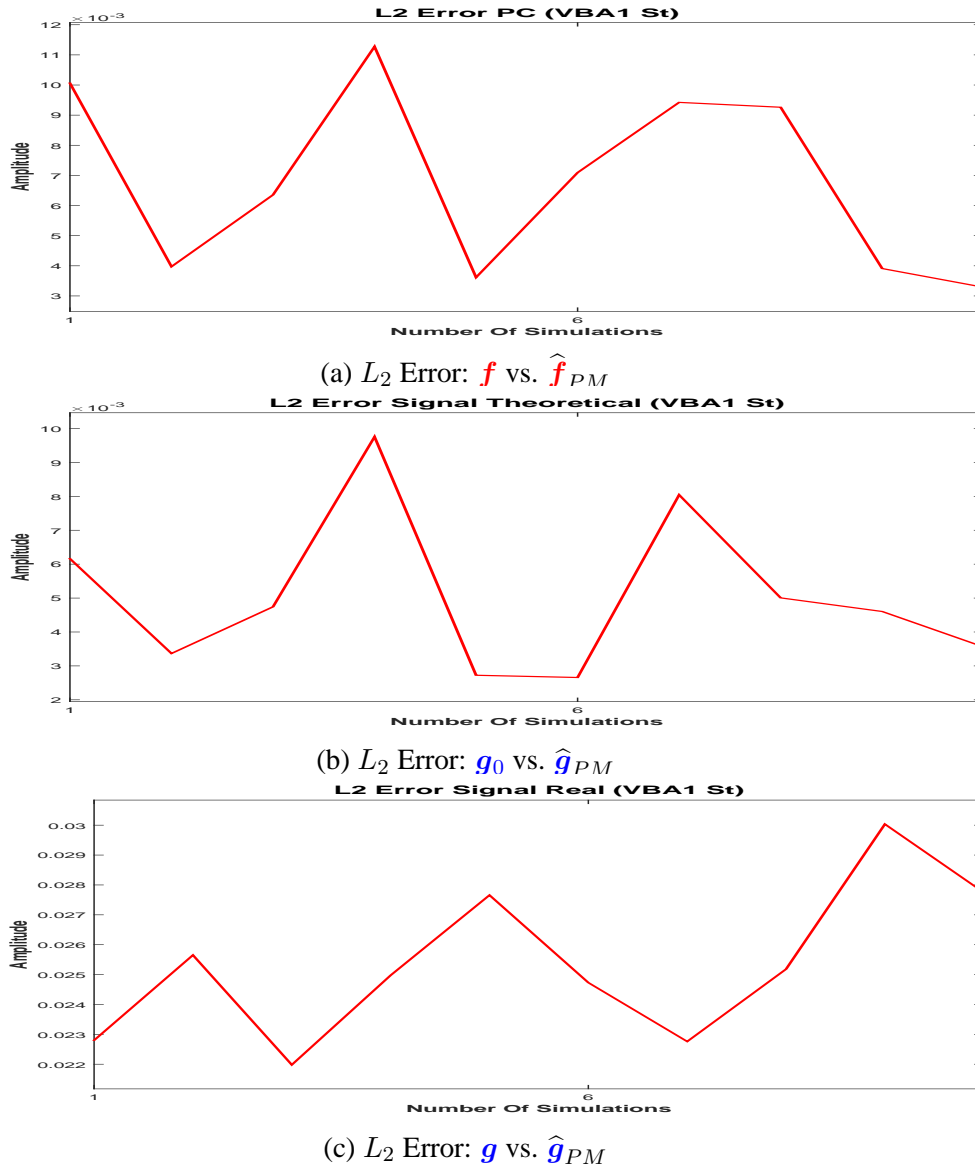
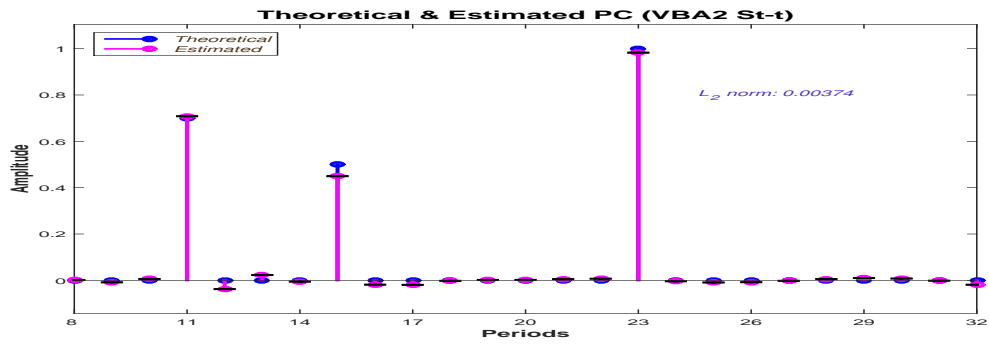


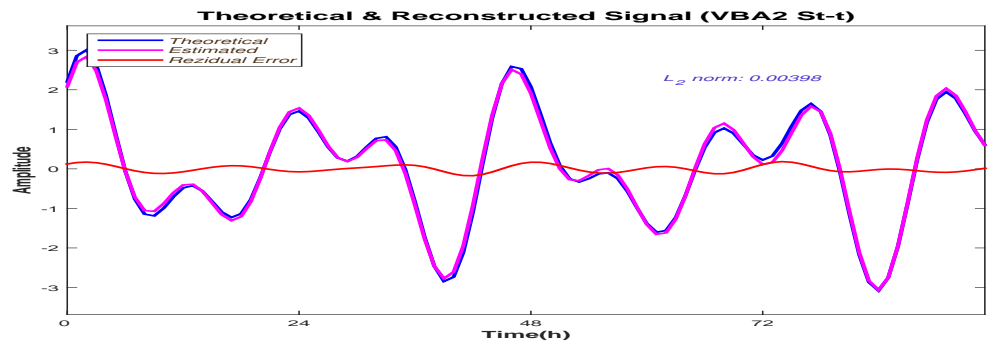
Figure 5.26: PM (via VBA, partial separability) IGSM  $L_2$  error measured for 10 different noise realisations (15dB)

### 5.3.4 PM (via VBA, full separability) IGSM 15dB

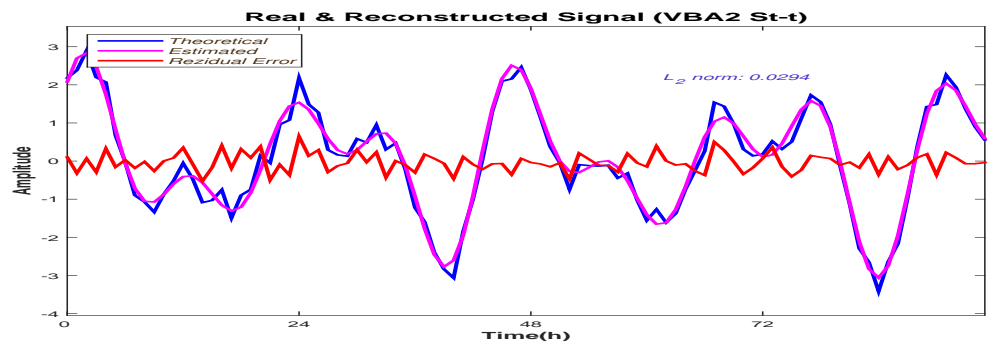
The estimations for the full separability case are also accurate: For the reconstruction of the theoretical signal  $g_0$ , the  $L_2$  error norm is  $\delta g_0 = \frac{\|g_0 - \hat{g}_{PM}\|_2^2}{\|g_0\|_2^2} = 0.00398$ . For the PC vector, the reconstruction error is  $\delta f = \frac{\|f - \hat{f}_{PM}\|_2^2}{\|f\|_2^2} = 0.00374$ . Figure (5.28a) presents the variation of  $L_2$  PC vector error reconstruction for 10 different noise realisation. The figure presents a very small variation of  $L_2$  PC vector error reconstruction, between 0.0037 and 0.027. Very small variations corresponding to the  $L_2$  error reconstruction for the theoretical signal  $g_0$ , and signal  $g$ , are presented in Figure (5.28b) and Figure (5.28c).



(a) Theoretical and Estimated PC



(b) Theoretical and Estimated Signal



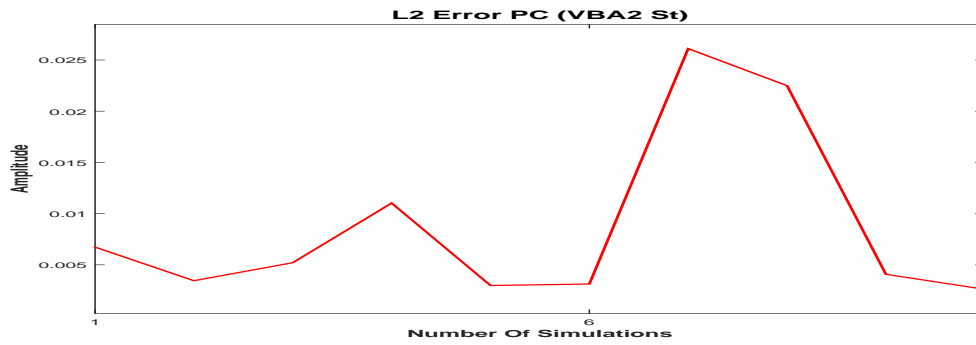
(c) Real and Estimated Signal

Figure 5.27: PM (via VBA, full separability) IGSM Estimation (15dB)

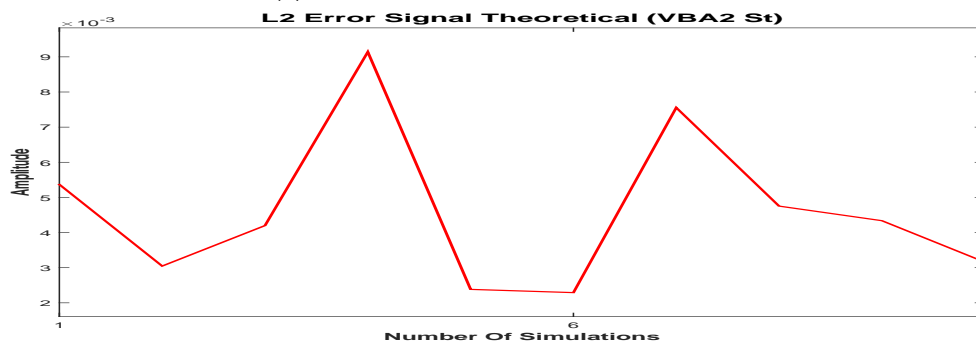
### 5.3.5 Methods comparison 15dB

The comparison between the estimations corresponding to the proposed IGSM models is presented in Figure (5.29c) (JMAP estimator), Figure (5.29d) (PM via VBA, partial separability estimator) and Figure (5.29e) (PM via VBA, full separability estimator). The comparison with the Gaussian case (i.e. Gaussian prior), via the two estimators discussed, is presented in Figure (5.29a) (Gaussian Model, JMAP estimator) and Figure (5.29b) (PM via VBA estimator). A comparison with the FFT is presented in Figure (5.29f). 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. VBA estimator, both partial separable and fully separable provides very accurate stable estimations.

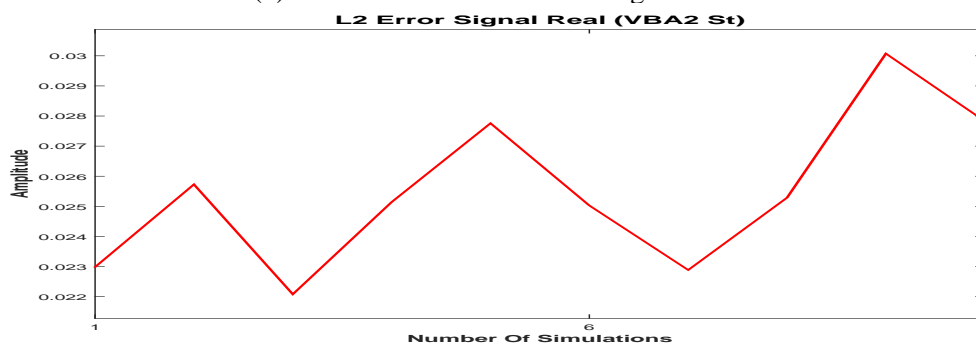
### 5.3.6 - Error comparison 15dB



(a) Theoretical and Estimated PC



(b) Theoretical and Estimated Signal



(c) Real and Estimated Signal

Figure 5.28: PM (via VBA, full separability) IGSM  $L_2$  error measured for 10 different noise realisations (15dB)

### 5.3.6 Error comparison 15dB

The  $L_2$  error measurement corresponding to the PC estimation, theoretical signal estimation and the signal estimation, for 10 different noise realisation is presented in the following figure:

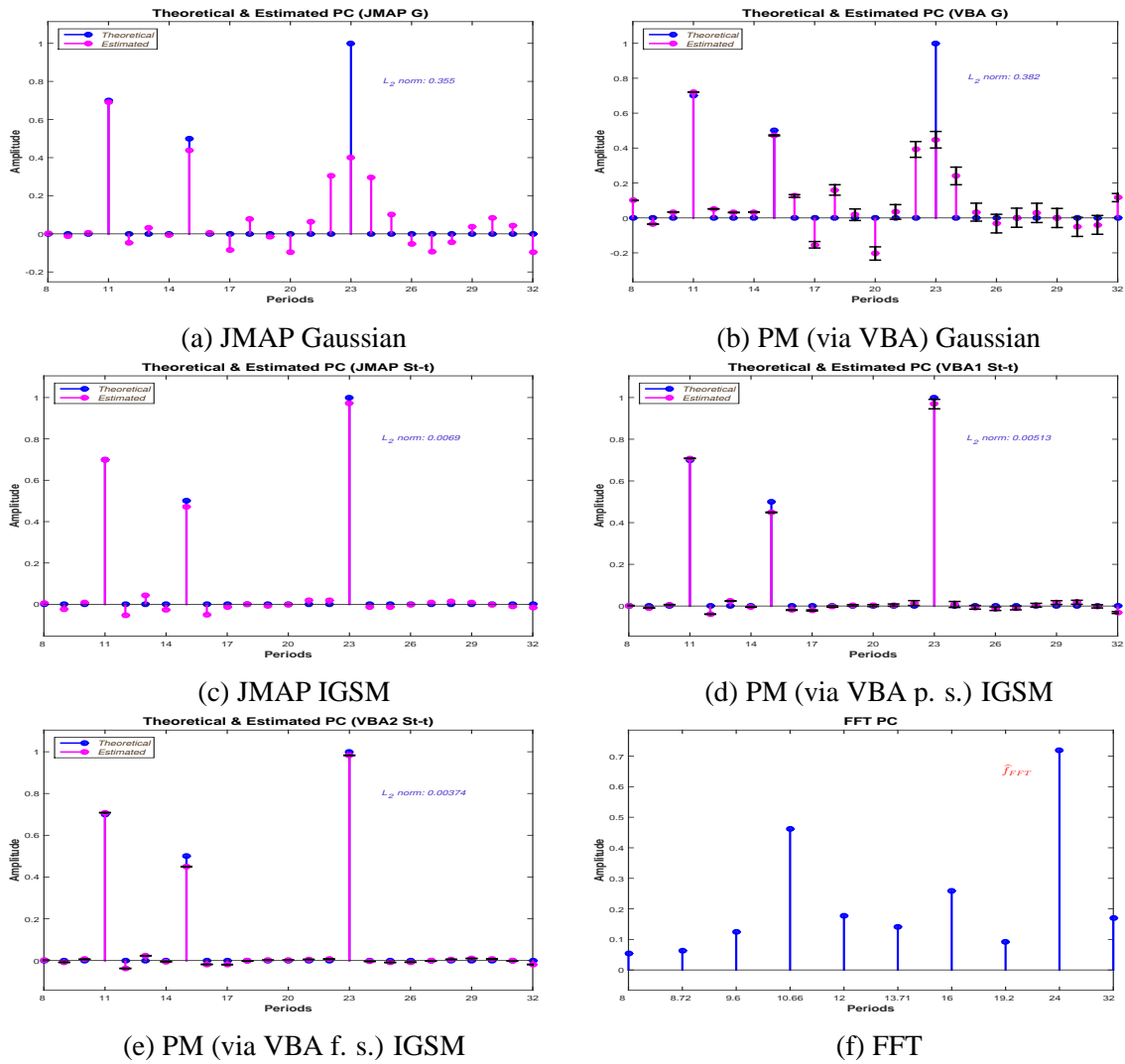
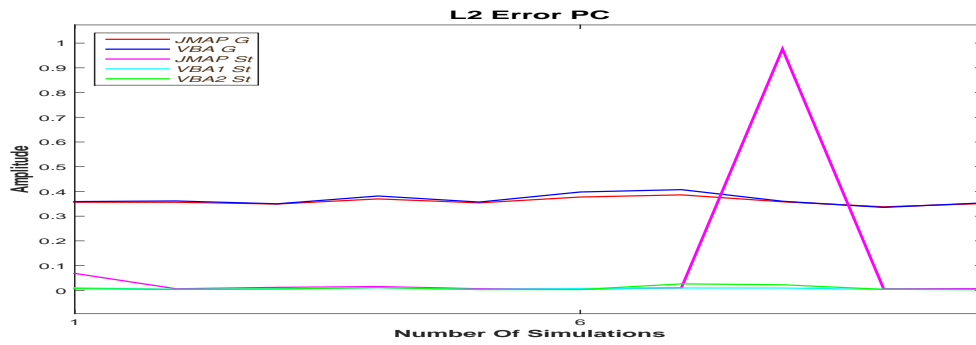
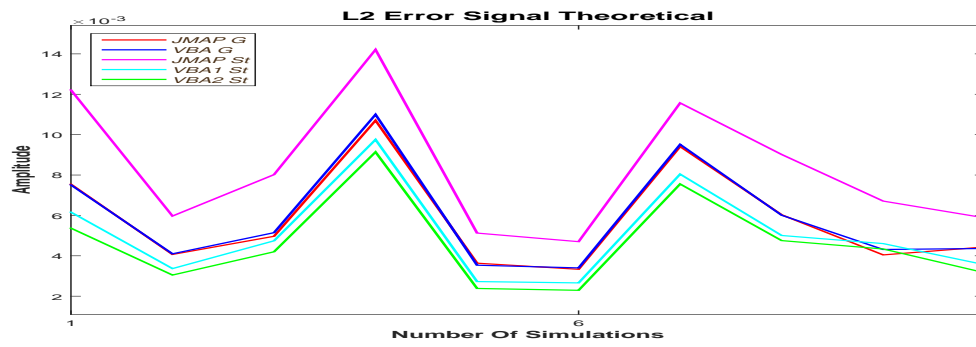


Figure 5.29: Methods Comparison (15dB)

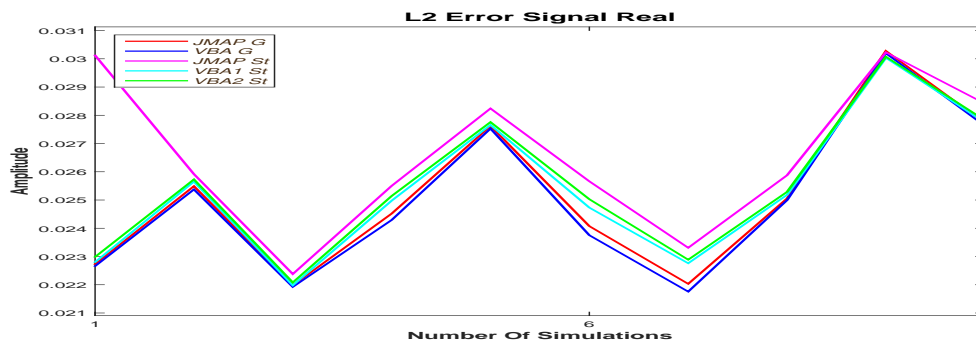
### 5.3.6 - Error comparison 15dB



(a) PC error estimation



(b) Theoretical signal error estimation



(c) Signal error estimation (15dB)

Figure 5.30:  $L_2$  Errors estimation (15dB)

## 5.4 Student-t PM via VBA vs. LASSO 05dB

In this section we compare the performances of the Proposed Method and LASSO, for SNR=05dB. We consider the theoretical PC amplitude vectors  $f_1$ ,  $f_2$  and  $f$ , defined as  $f = \sqrt{f_1^2 + f_2^2}$ , the corresponding theoretical (original) signal  $g_0 = H_1 f_1 + H_2 f_2$  and the corresponding (noisy) signal  $g = g_0 + \epsilon$ , by adding noise over the theoretical signal. The considered signal represents a four day signal, sampled every hour. The matrix  $H_1$  is a sine matrix and the matrix  $H_2$  is a cosine matrix. The data is presented in Figure (5.31):

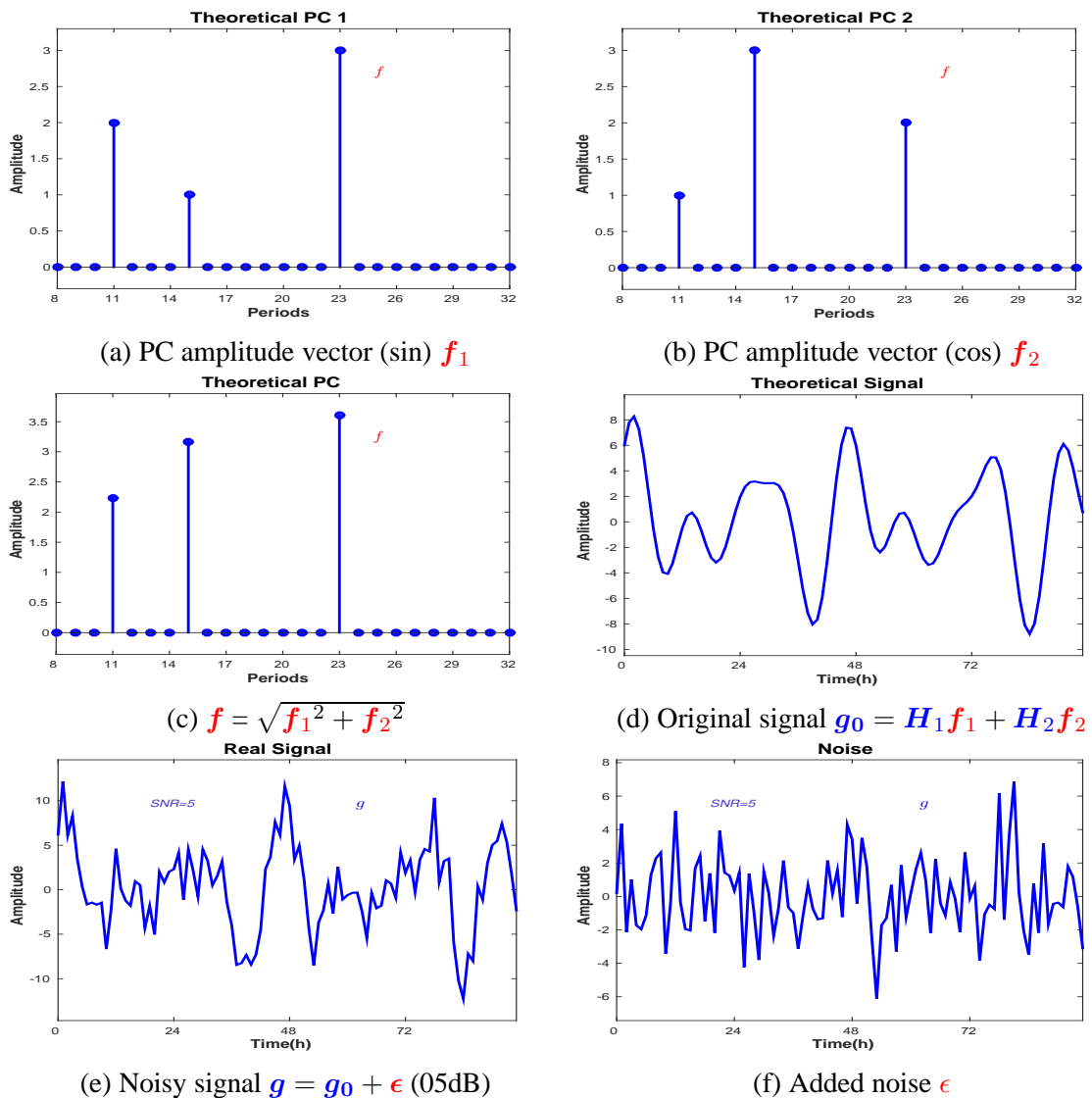


Figure 5.31: Synthetic simulation data (05dB): sine PC amplitudes vector  $f_1$ , (5.31a), cosine PC amplitudes vector  $f_2$ , (5.31b), PC amplitudes vector  $f$ , (5.31c); corresponding original signal  $g_0 = H_1 f_1 + H_2 f_2$ , (5.31d), noisy signal  $g$ , (5.31e) and the added noise (05dB)  $\epsilon$  (5.31f)

the PC amplitude vector  $f_1$ , corresponding to sine, is presented in Figure (5.31a) and the PC amplitude vector  $f_2$ , corresponding to cosine, is presented in Figure (5.31b). We have considered sparse vectors, with three periods with non-zero corresponding amplitudes: 11 hours, 15 hours and 23 hours. We have considered for both PC amplitudes vectors  $f_1$  and  $f_2$  the same

### 5.3.6 - Error comparison 15dB

periods with non-zero corresponding amplitudes, but different values for the amplitudes. The PC amplitude vector  $\mathbf{f}$  is presented in Figure (5.31c), the corresponding original signal  $\mathbf{g}_0$  in Figure (5.31d) and the corresponding noisy signal  $\mathbf{g}$  in Figure (5.31e). The added noise  $\boldsymbol{\epsilon}$ , corresponding to SNR=05dB, is presented in Figure (5.31f). The estimation corresponding to the Student-t model, with PM estimation, is presented in Figure (5.32). The comparison between the estimated  $\hat{\mathbf{f}}_1$  and  $\mathbf{f}_1$  is presented in Figure (5.32a) and the comparison between the estimated  $\hat{\mathbf{f}}_2$  and  $\mathbf{f}_2$  is presented in Figure (5.32b). Both estimated PC amplitudes vectors  $\hat{\mathbf{f}}_1$  and  $\hat{\mathbf{f}}_2$  are sparse and the non-zero amplitudes are correctly estimated. The  $L_1$  and  $L_2$  estimation error norms corresponding to  $\mathbf{f}_1$  are  $\delta \mathbf{f}_{11} = 0.34$  and  $\delta \mathbf{f}_{12} = 0.05$ ; the  $L_1$  and  $L_2$  estimation error norms corresponding to  $\mathbf{f}_2$  are  $\delta \mathbf{f}_{21} = 0.16$  and  $\delta \mathbf{f}_{22} = 0.004$ . The

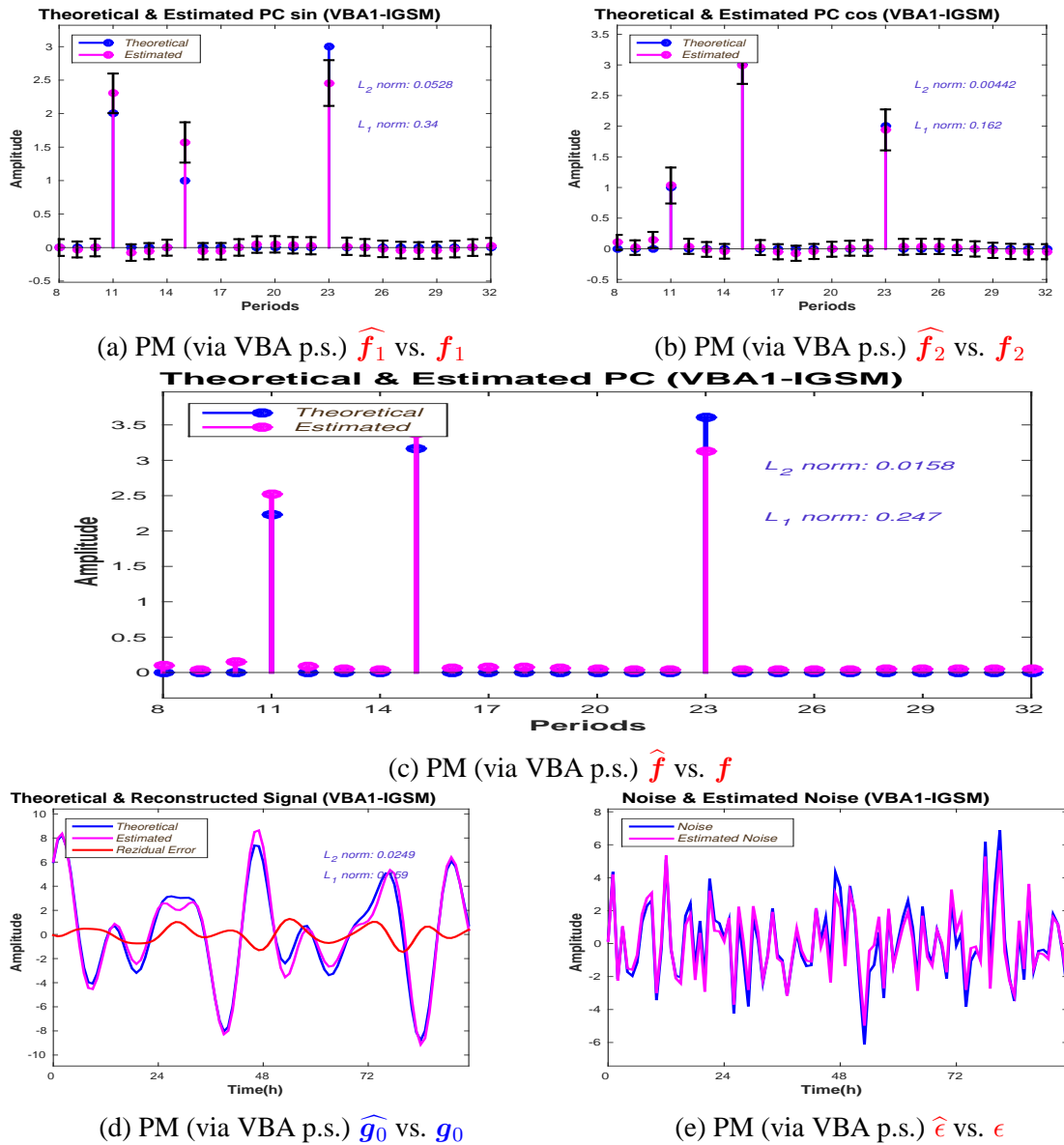


Figure 5.32: Synthetic simulation (05dB): comparison between the Proposed Method estimations and the synthetic data

comparison between  $\mathbf{f}$  and  $\hat{\mathbf{f}}$  is presented in Figure (5.32c): the  $L_1$  estimation error norm is  $\delta \mathbf{f}_1 = 0.247$  and the  $L_2$  estimation error norm is  $\delta \mathbf{f}_2 = 0.015$ . The comparison between the reconstructed signal  $\hat{\mathbf{g}}$  and  $\mathbf{g}$  is presented in Figure (5.32d): the reconstruction is fairly

accurate: the  $L_1$  reconstruction error norm is  $\delta g_1 = 0.159$  and the  $L_2$  reconstruction error norm is  $\delta g_2 = 0.024$ . The estimation corresponding to LASSO is presented in Figure (5.33).

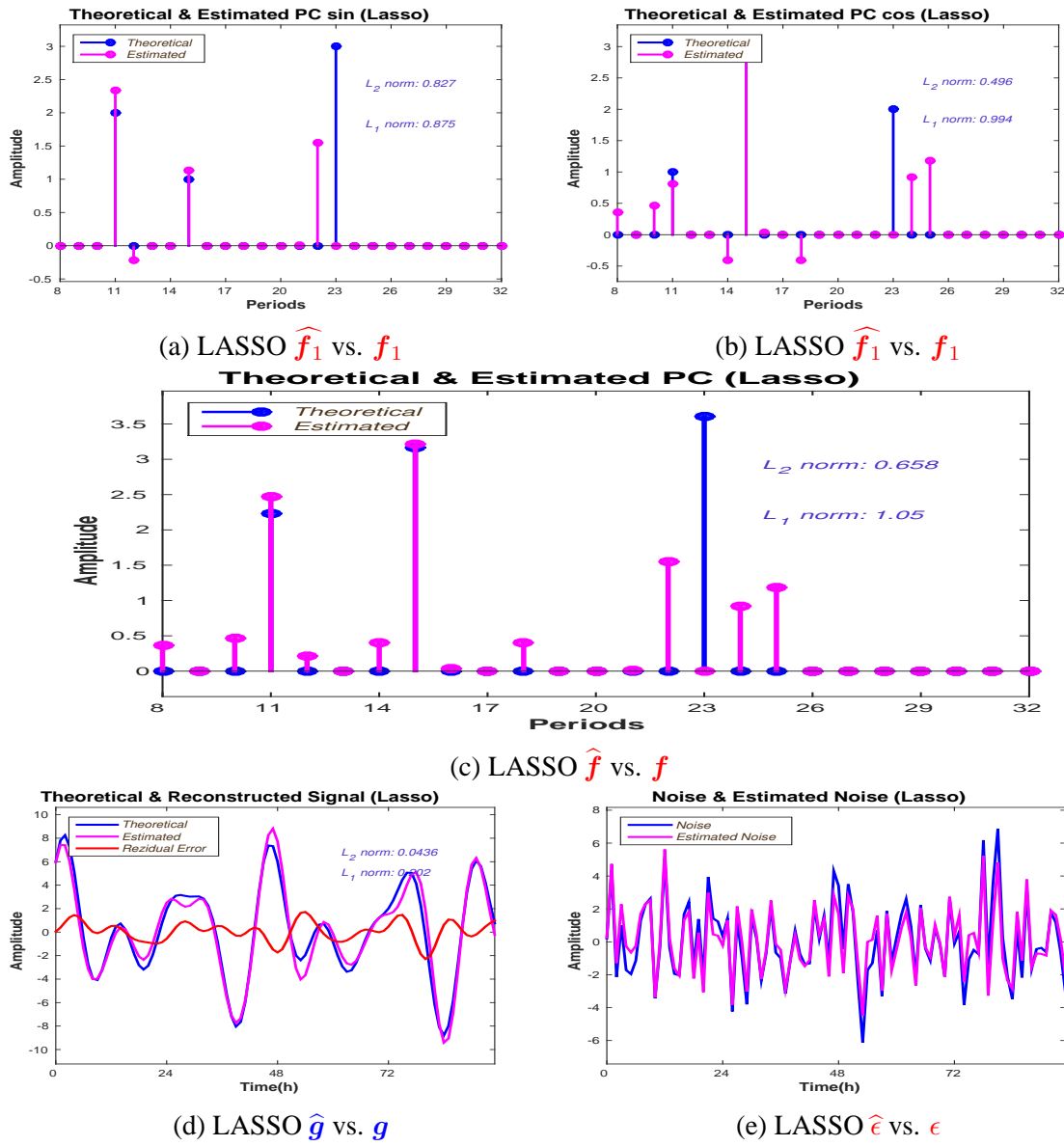


Figure 5.33: Synthetic simulation (05dB): comparison between the LASSO estimations and the synthetic data

The comparison between the estimated  $\hat{f}_1$  and  $f_1$  is presented in Figure (5.33a) and the comparison between the estimated  $\hat{f}_2$  and  $f_2$  is presented in Figure (5.33b). Both estimations have a *missing value* (i.e. the non-zero amplitude corresponding to 23h hours is estimated as zero) and both estimations have one or two *false detections*: for  $\hat{f}_1$ , LASSO is associating a non-zero amplitude corresponding to 22 hours and for  $\hat{f}_2$  two non-zero amplitudes corresponding to 24 and 25 hours. The  $L_1$  and  $L_2$  estimation error norms corresponding to  $f_1$  are  $\delta f_{11} = 0.87$  and  $\delta f_{12} = 0.82$  and the  $L_1$  and  $L_2$  estimation error norms corresponding to  $f_2$  are  $\delta f_{21} = 0.49$  and  $\delta f_{22} = 0.99$ . The comparison between  $f$  and  $\hat{f}$  is presented in Figure (5.33c): the  $L_2$  estimation error norm is  $\delta f_2 = 0.65$  and the  $L_1$  estimation error norm is  $\delta f_1 = 1.05$ . The comparison between the reconstructed signal  $\hat{g}$  and  $g$  is presented in Figure (5.33d): the reconstruction is fairly accurate: the  $L_1$  reconstruction error norm is



### 5.3.6 - Error comparison 15dB

$\delta g_1 = 0.2$  and the  $L_2$  reconstruction error norm is  $\delta g_2 = 0.04$ . To analyse the stability of

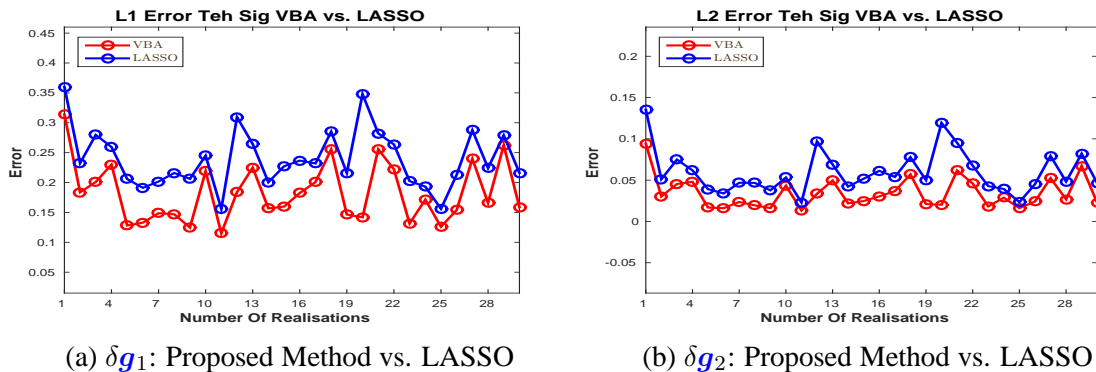


Figure 5.34: Stability: L1 and L2 reconstruction errors (05dB):  $\delta g_1$  and  $\delta g_2$ : Proposed Method vs. LASSO

the Proposed Method and LASSO, we compute the  $L_1$  and  $L_2$  estimation and reconstruction errors corresponding to the Proposed Method and LASSO, for 30 noise realisations, in the same conditions as in Figure (5.31). Figure (5.34) presents the behaviour of the  $L_1$  and  $L_2$  reconstruction errors corresponding to the original signal  $g_0$ : Figure (5.34a) compares the behaviour of  $L_1$  reconstruction error corresponding to the original signal,  $\delta g_1$ , corresponding to the Proposed Method and LASSO and Figure (5.34b) compares the behaviour of  $L_2$  reconstruction error corresponding to the original signal,  $\delta g_2$ , corresponding to the Proposed Method and LASSO. Both methods provide small reconstruction errors, and both methods are stable. Figure (5.35) presents the behaviour of the  $L_1$  and  $L_2$  estimation errors for  $f_1$ ,  $f_2$  and  $f$ : Figure (5.35a) compares the behaviour of  $L_1$  estimation error for  $f_1$ ,  $\delta f_{11}$ , corresponding to the Proposed Method and LASSO and Figure (5.35b) compares the behaviour of  $L_2$  estimation error for  $f_2$ ,  $\delta f_{12}$ , corresponding to the Proposed Method and LASSO. The comparison corresponding to  $f_2$  is presented in Figure (5.35c) ( $L_1$ ) and Figure (5.35d) ( $L_2$ ) and corresponding to  $f$  is presented in Figure (5.35e) ( $L_1$ ) and Figure (5.35f) ( $L_2$ ).

The Proposed Method has better performances compared to LASSO in terms of  $L_1$  and  $L_2$  estimation errors for the PC amplitudes vector: the associated errors are always smaller than the ones corresponding to LASSO. Also, in this case, for SNR=05dB the LASSO is unstable.

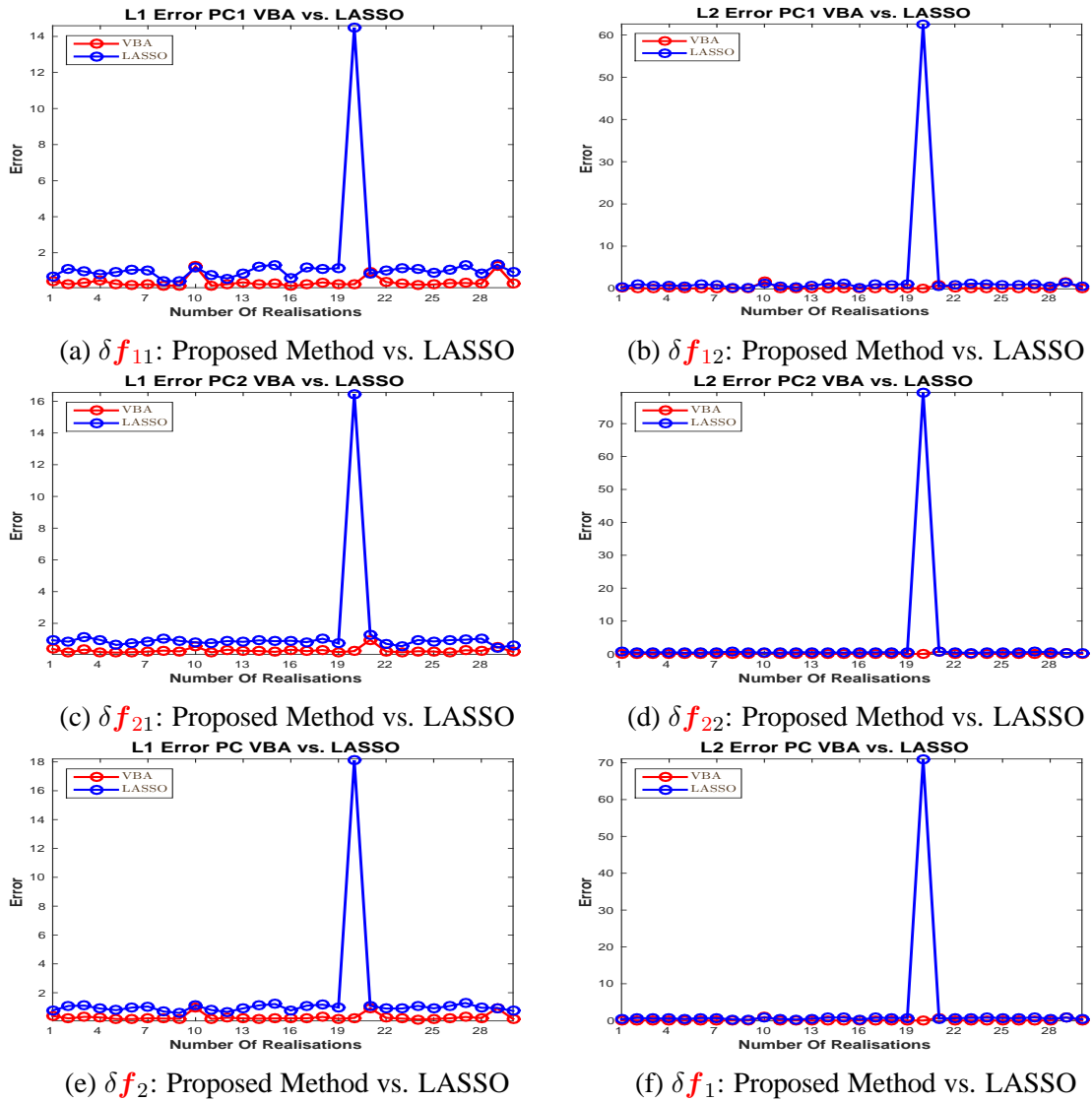


Figure 5.35: Stability: L1 and L2 estimation errors (05dB):  $\delta f_1$  and  $\delta f_2$ : Proposed Method vs. LASSO

## 5.5 Student-t PM via VBA vs. LASSO 10dB

In this section we compare the performances of the Proposed Method and LASSO, for SNR=10dB. We consider the theoretical PC amplitude vectors  $f_1$ ,  $f_2$  and  $f$ , defined as  $f = \sqrt{f_1^2 + f_2^2}$ , the corresponding theoretical (original) signal  $g_0 = H_1 f_1 + H_2 f_2$  and the corresponding (noisy) signal  $g = g_0 + \epsilon$ , by adding noise over the theoretical signal. The considered signal represents a four day signal, sampled every hour. The matrix  $H_1$  is a sine matrix and the matrix  $H_2$  is a cosine matrix. The data is presented in Figure (5.36):

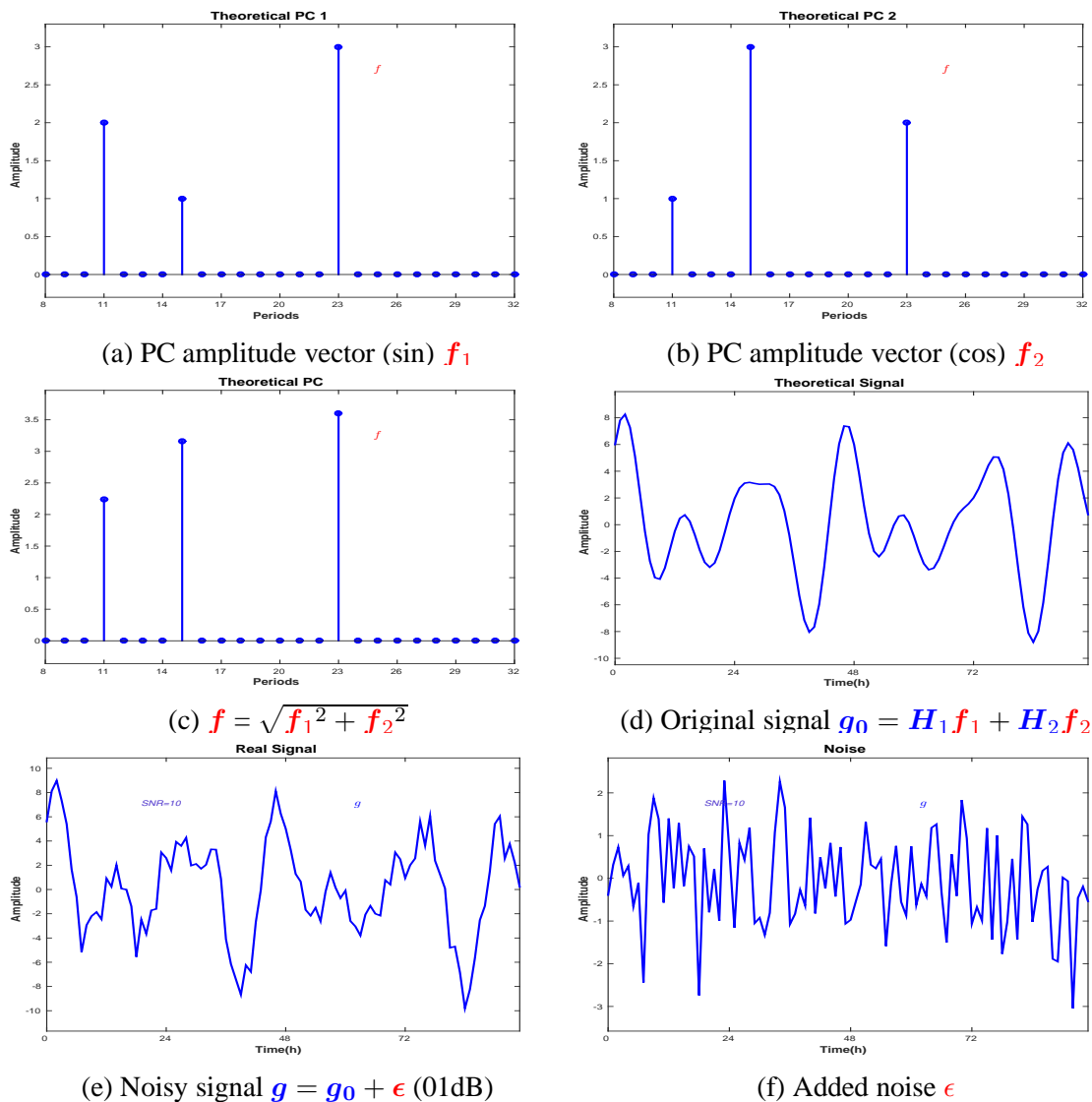


Figure 5.36: Synthetic simulation data (10dB): sine PC amplitudes vector  $f_1$ , (5.36a), cosine PC amplitudes vector  $f_2$ , (5.36b), PC amplitudes vector  $f$ , (5.36c); corresponding original signal  $g_0 = H_1 f_1 + H_2 f_2$ , (5.36d), noisy signal  $g$ , (5.36e) and the added noise (05dB)  $\epsilon$  (5.36f)

the PC amplitude vector  $f_1$ , corresponding to sine, is presented in Figure (5.36a) and the PC amplitude vector  $f_2$ , corresponding to cosine, is presented in Figure (5.36b). We have considered sparse vectors, with three periods with non-zero corresponding amplitudes: 11 hours, 15 hours and 23 hours. We have considered for both PC amplitudes vectors  $f_1$  and  $f_2$  the same

periods with non-zero corresponding amplitudes, but different values for the amplitudes. The PC amplitude vector  $\mathbf{f}$  is presented in Figure (5.36c), the corresponding original signal  $\mathbf{g}_0$  in Figure (5.36d) and the corresponding noisy signal  $\mathbf{g}$  in Figure (5.36e). The added noise  $\epsilon$ , corresponding to SNR=10dB, is presented in Figure (5.36f). The estimation corresponding

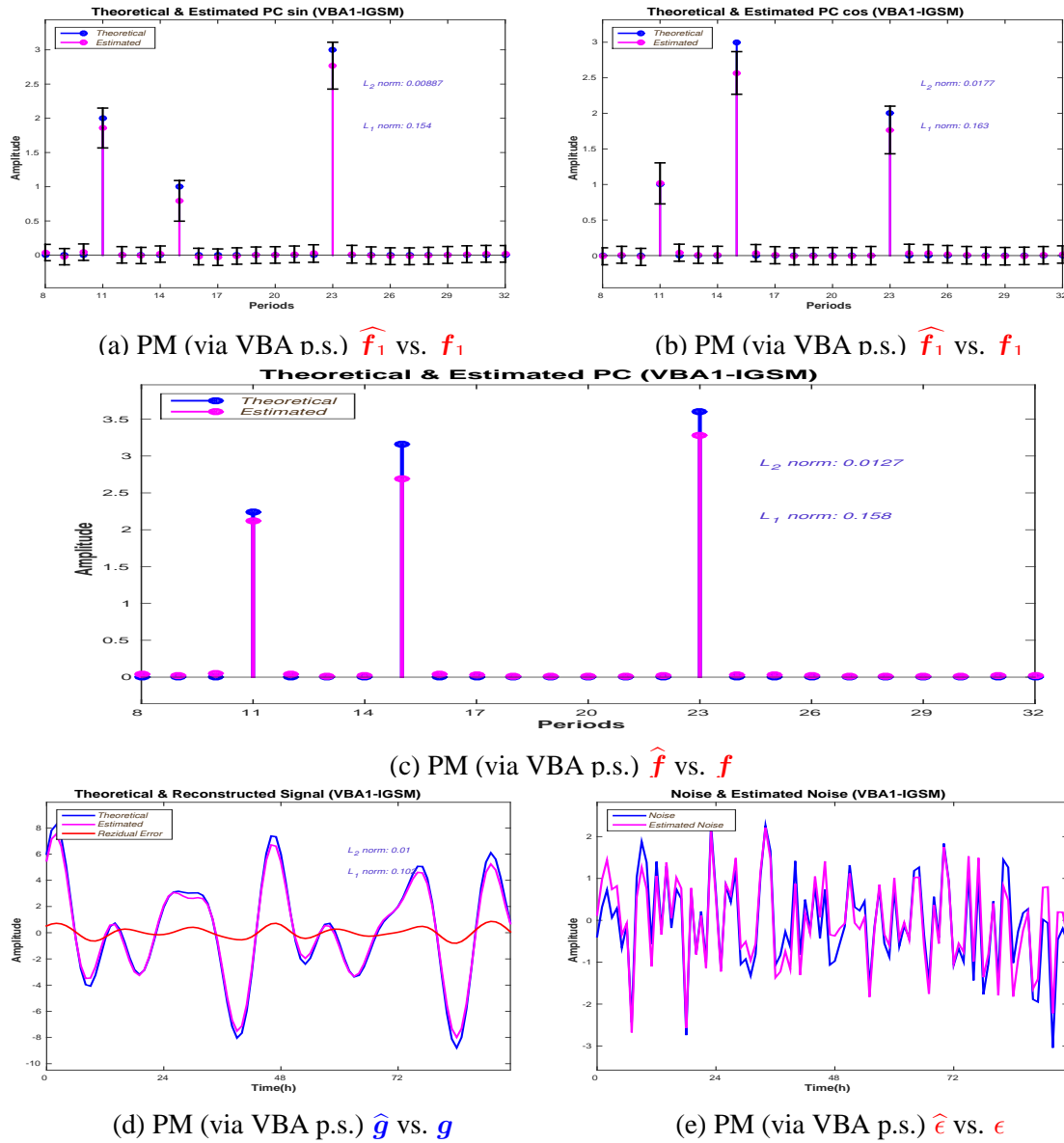


Figure 5.37: Synthetic simulation (10dB): comparison between the Proposed Method estimations and the synthetic data

to the Student-t model, with PM estimation, is presented in Figure (5.37). The comparison between the estimated  $\hat{\mathbf{f}}_1$  and  $\mathbf{f}_1$  is presented in Figure (5.37a) and the comparison between the estimated  $\hat{\mathbf{f}}_2$  and  $\mathbf{f}_2$  is presented in Figure (5.37b). Both estimated PC amplitudes vectors  $\hat{\mathbf{f}}_1$  and  $\hat{\mathbf{f}}_2$  are sparse and the non-zero amplitudes are correctly estimated. The  $L_1$  and  $L_2$  estimation error norms corresponding to  $\mathbf{f}_1$  are  $\delta \mathbf{f}_{11} = 0.15$  and  $\delta \mathbf{f}_{12} = 0.008$ ; the  $L_1$  and  $L_2$  estimation error norms corresponding to  $\mathbf{f}_2$  are  $\delta \mathbf{f}_{21} = 0.16$  and  $\delta \mathbf{f}_{22} = 0.017$ .

The comparison between  $\mathbf{f}$  and  $\hat{\mathbf{f}}$  is presented in Figure (5.37c): the  $L_1$  estimation error norm is  $\delta \mathbf{f}_1 = 0.158$  and the  $L_2$  estimation error norm is  $\delta \mathbf{f}_2 = 0.012$ . The comparison between the reconstructed signal  $\hat{\mathbf{g}}$  and  $\mathbf{g}$  is presented in Figure (5.37d): the reconstruction is

### 5.3.6 - Error comparison 15dB

fairly accurate: the  $L_1$  reconstruction error norm is  $\delta g_1 = 0.102$  and the  $L_2$  reconstruction error norm is  $\delta g_2 = 0.01$ . The estimation corresponding to LASSO is presented in Figure (5.38).

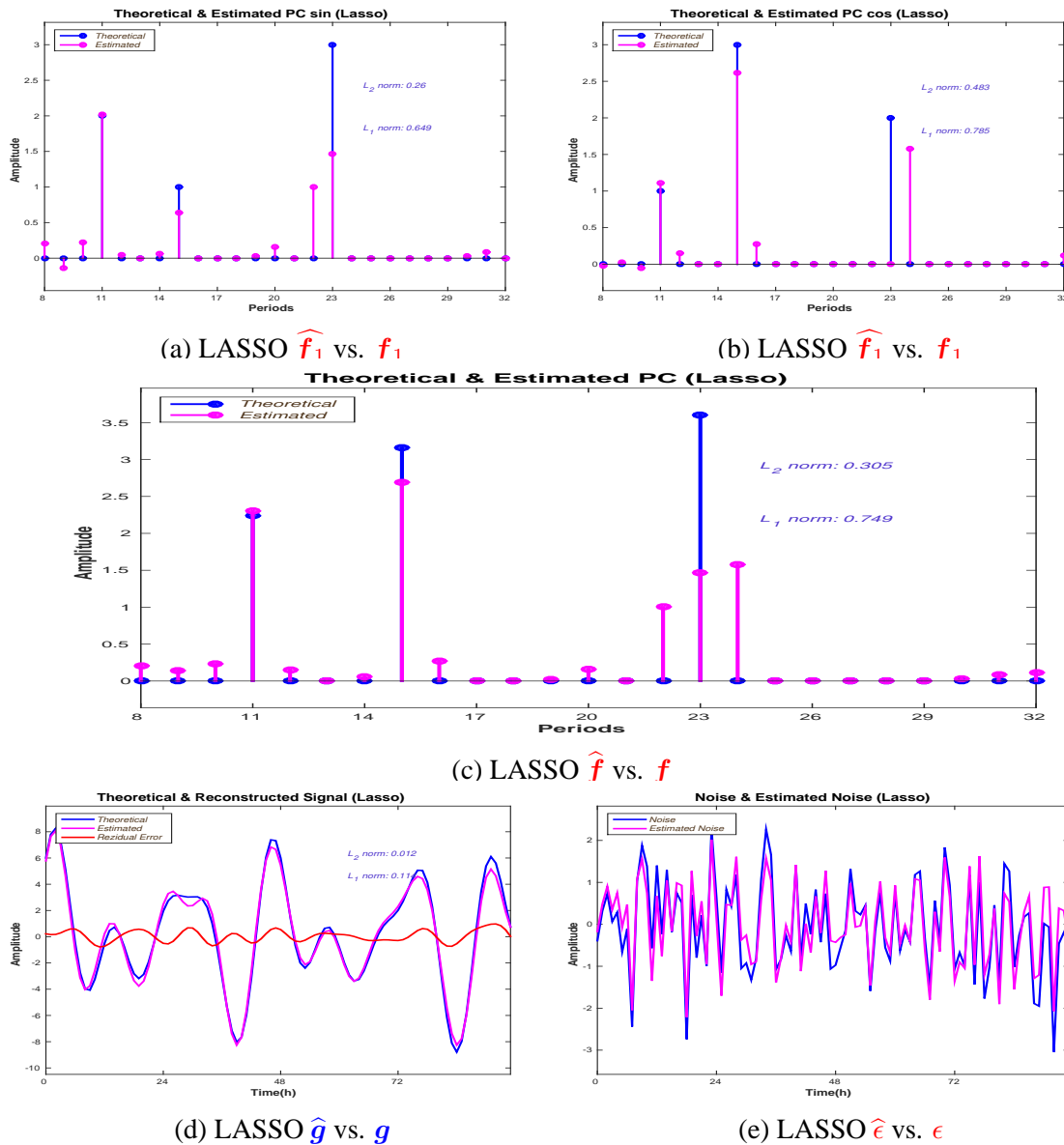


Figure 5.38: Synthetic simulation (10dB): comparison between the LASSO estimations and the synthetic data

The comparison between the estimated  $\hat{f}_1$  and  $f_1$  is presented in Figure (5.38a) and the comparison between the estimated  $\hat{f}_2$  and  $f_2$  is presented in Figure (5.38b). Both estimations have a *missing value* (i.e. the non-zero amplitude corresponding to 23h hours is estimated as zero) and both estimations have one or two *false detections*: for  $\hat{f}_1$ , LASSO is associating a non-zero amplitude corresponding to 22 hours and for  $\hat{f}_2$  two non-zero amplitudes corresponding to 24 and 25 hours. The  $L_1$  and  $L_2$  estimation error norms corresponding to  $f_1$  are  $\delta f_{11} = 0.649$  and  $\delta f_{12} = 0.26$  and the  $L_1$  and  $L_2$  estimation error norms corresponding to  $f_2$  are  $\delta f_{21} = 0.78$  and  $\delta f_{22} = 0.48$ . The comparison between  $f$  and  $\hat{f}$  is presented in Figure (5.38c): the  $L_2$  estimation error norm is  $\delta f_2 = 0.305$  and the  $L_1$  estimation error norm is  $\delta f_1 = 0.749$ . The comparison between the reconstructed signal  $\hat{g}$  and  $g$  is presented in Figure (5.38d): the reconstruction is fairly accurate: the  $L_1$  reconstruction error norm is

$\delta g_1 = 0.012$  and the  $L_2$  reconstruction error norm is  $\delta g_2 = 0.11$ . To analyse the stability

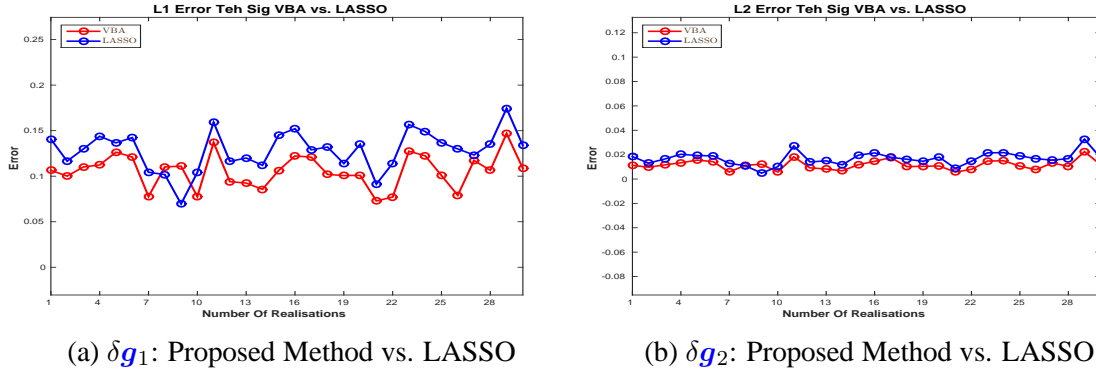


Figure 5.39: Stability:  $L_1$  and  $L_2$  reconstruction errors (10dB):  $\delta g_1$  and  $\delta g_2$ : Proposed Method vs. LASSO

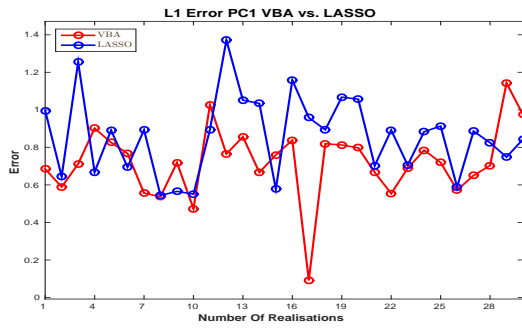
of the Proposed Method and LASSO, we compute the  $L_1$  and  $L_2$  estimation and reconstruction errors corresponding to the Proposed Method and LASSO, for 30 noise realisations, in the same conditions as in Figure (5.36). Figure (5.39) presents the behaviour of the  $L_1$  and  $L_2$  reconstruction errors corresponding to the original signal  $g_0$ : Figure (5.39a) compares the behaviour of  $L_1$  reconstruction error corresponding to the original signal,  $\delta g_1$ , corresponding to the Proposed Method and LASSO and Figure (5.39b) compares the behaviour of  $L_2$  reconstruction error corresponding to the original signal,  $\delta g_2$ , corresponding to the Proposed Method and LASSO. Both methods provide small reconstruction errors, and both methods are stable. Figure (5.40) presents the behaviour of the  $L_1$  and  $L_2$  estimation errors for  $f_1$ ,  $f_2$  and  $f$ : Figure (5.40a) compares the behaviour of  $L_1$  estimation error for  $f_1$ ,  $\delta f_{11}$ , corresponding to the Proposed Method and LASSO and Figure (5.40b) compares the behaviour of  $L_2$  estimation error for  $f_2$ ,  $\delta f_{12}$ , corresponding to the Proposed Method and LASSO. The comparison corresponding to  $f_2$  is presented in Figure (5.40c) ( $L_1$ ) and Figure (5.40d) ( $L_2$ ) and corresponding to  $f$  is presented in Figure (5.40e) ( $L_1$ ) and Figure (5.40f) ( $L_2$ )

The Proposed Method has better performances compared to LASSO in terms of  $L_1$  and  $L_2$  estimation errors for the PC amplitudes vector: the associated errors are always smaller than the ones corresponding to LASSO.

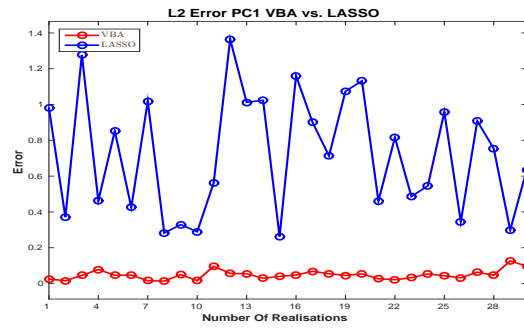
Finally, we present a comparison between the Proposed Method and LASSO for different values for the SNR. Figure (5.41a) presents the comparison between the  $L_1$  reconstruction norm for  $g_0$  corresponding to the Proposed Method and LASSO for  $\text{SNR}=\{5\text{dB}, 10\text{dB}, 15\text{dB}, 20\text{dB}, 25\text{dB}, 30\text{dB}, 40\text{dB}, 50\text{dB}\}$ . The comparison between the  $L_2$  reconstruction norm for  $g_0$  for the two methods is presented in Figure (5.41b). Figure (5.41c) and Figure (5.41d) present the comparison between the  $L_1$  and  $L_2$  reconstruction norm for  $g$  for the two methods.

Figure (5.42a) presents the comparison between the  $L_1$  estimation norm for  $f_1$  corresponding to the Proposed Method and LASSO for  $\text{SNR}=\{5\text{dB}, 10\text{dB}, 15\text{dB}, 20\text{dB}, 25\text{dB}, 30\text{dB}, 40\text{dB}, 50\text{dB}\}$ . The comparison between the  $L_2$  reconstruction norm for  $f_1$  for the two methods is presented in Figure (5.42b). Figure (5.42c) and Figure (5.42d) present the comparison between the  $L_1$  and  $L_2$  estimation error norm for  $f_2$  for the two methods and Figure (5.42e) and Figure (5.42f) present the comparison between the  $L_1$  and  $L_2$  estimation error norm for  $f$  for the two methods.

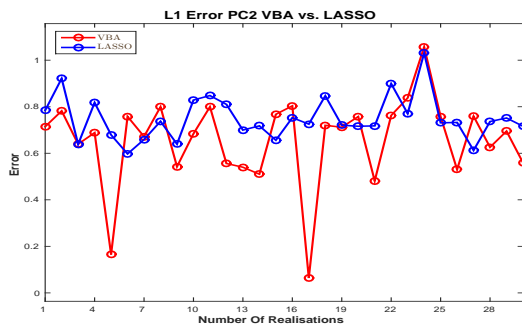
### 5.3.6 - Error comparison 15dB



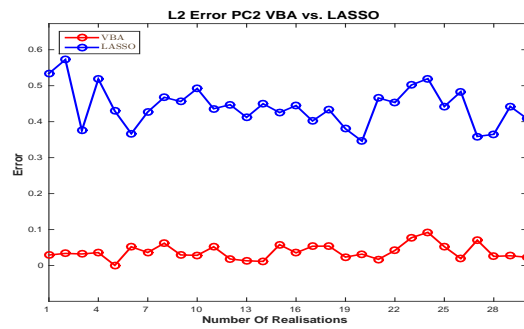
(a)  $\delta f_{11}$ : Proposed Method vs. LASSO



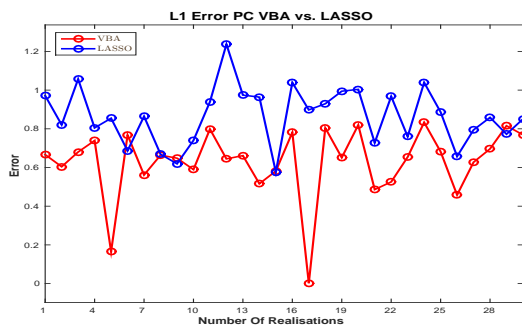
(b)  $\delta f_{12}$ : Proposed Method vs. LASSO



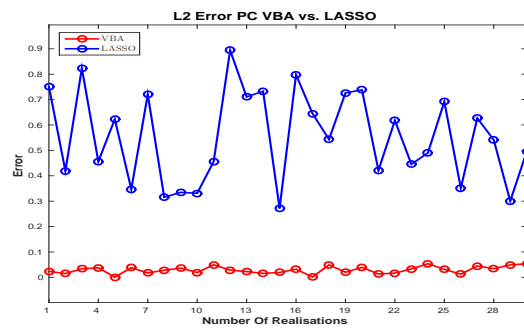
(c)  $\delta f_{21}$ : Proposed Method vs. LASSO



(d)  $\delta f_{22}$ : Proposed Method vs. LASSO



(e)  $\delta f_2$ : Proposed Method vs. LASSO



(f)  $\delta f_1$ : Proposed Method vs. LASSO

Figure 5.40: Stability: L1 and L2 estimation errors (10dB):  $\delta f_1$  and  $\delta f_2$ : Proposed Method vs. LASSO

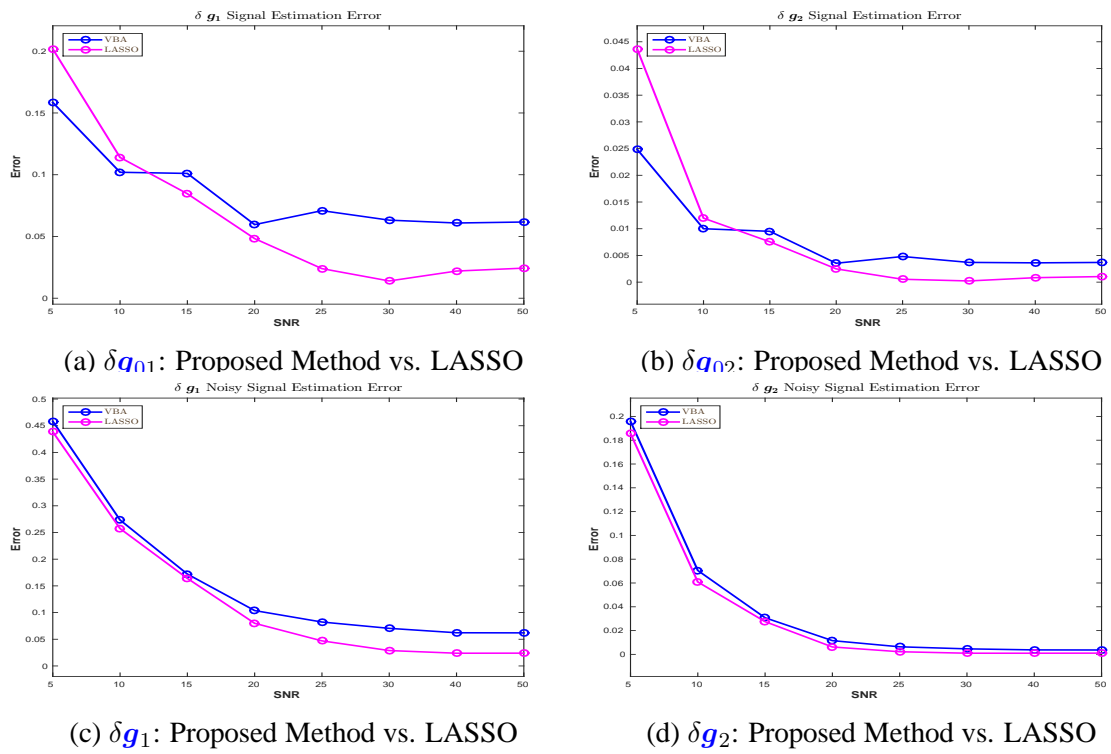
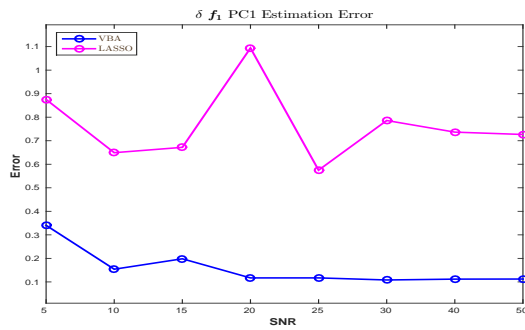


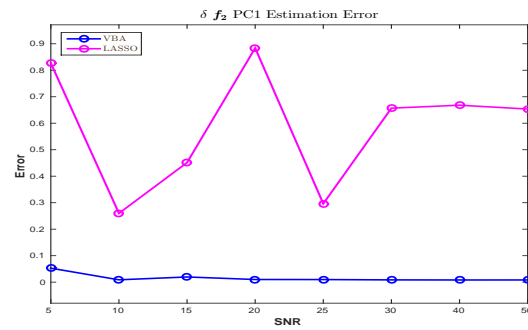
Figure 5.41: L1 and L2 estimation errors:  $\delta g_1$  and  $\delta g_2$ : Proposed Method vs. LASSO



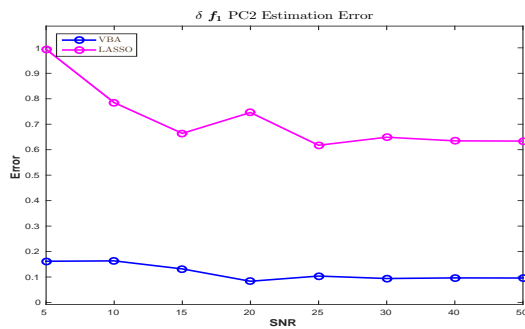
### 5.3.6 - Error comparison 15dB



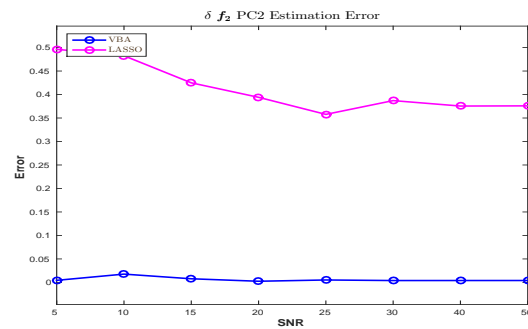
(a)  $\delta f_{11}$ : Proposed Method vs. LASSO



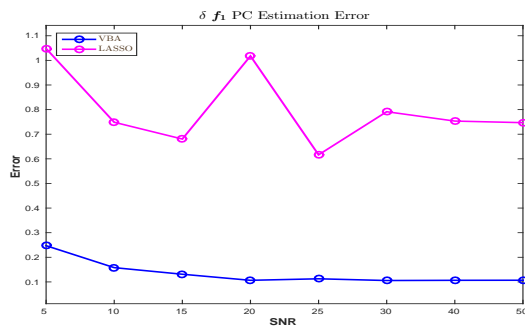
(b)  $\delta f_{12}$ : Proposed Method vs. LASSO



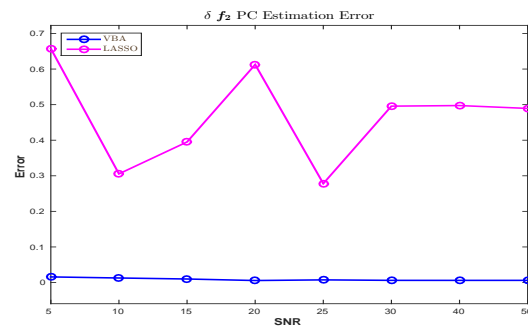
(c)  $\delta f_{11}$ : Proposed Method vs. LASSO



(d)  $\delta f_{12}$ : Proposed Method vs. LASSO



(e)  $\delta f_{21}$ : Proposed Method vs. LASSO



(f)  $\delta f_{22}$ : Proposed Method vs. LASSO

Figure 5.42: L1 and L2 estimation errors:  $\delta f_1$  and  $\delta f_2$ : Proposed Method vs. LASSO



## Simulations: Real data

### Contents

---

<b>6.1</b>	<b>CT 502 A1: Photon Absorption and Activity</b>	<b>122</b>
6.1.1	Dominant Period evolution: Activity	123
6.1.2	Dominant Period evolution: Photon	127
6.1.3	Conclusions: Dominant Period evolution	131
<b>6.2</b>	<b>CT 502 A2: Photon Absorption and Activity</b>	<b>132</b>
6.2.1	Dominant Period evolution: Activity	132
6.2.2	Dominant Period evolution: Photon	137
6.2.3	Conclusions: Dominant Period evolution	140
<b>6.3</b>	<b>CT 502 A3: Photon Absorption and Activity</b>	<b>142</b>
6.3.1	Dominant Period evolution: Activity	142
6.3.2	Dominant Period evolution: Photon	146
6.3.3	Conclusions: Dominant Period evolution	150
<b>6.4</b>	<b>CT 502 B1: Photon Absorption and Activity</b>	<b>151</b>
6.4.1	Dominant Period evolution: Activity	152
6.4.2	Dominant Period evolution: Photon	156
6.4.3	Conclusions: Dominant Period evolution	161
<b>6.5</b>	<b>CT 502 B2: Photon Absorption and Activity</b>	<b>161</b>
6.5.1	Dominant Period evolution: Activity	162
6.5.2	Dominant Period evolution: Photon	164
<b>6.6</b>	<b>CT 502 B3: Photon Absorption and Activity</b>	<b>168</b>
6.6.1	Dominant Period evolution: Activity	169
6.6.2	Dominant Period evolution: Photon	172

**I**N this chapter we present the simulations corresponding to the real data. The considered signals are obtained in an experiment performed in 6 male KI/KI Per2::luc mice, aged 10 weeks. Mice were singly housed in RT-BIO and synchronized with LD12:12 for 10 days of adaptation. D-luciferin ([1.5 mg/mL]) was loaded in drink water. Animals were then kept in DD for 3 days. Vehicle or everolimus (5 mg/kg/d) was daily administered at ZT1 or ZT13 (Zeitgeber time) for 6 days, and then animals were kept in DD for another 4 days. Liver bioluminescence and mouse rest-activity were monitored every minute during LD12:12 and/or DD with the RT-BIO photomultiplier tube and infrared sensor respectively. For each mouse, CT 502 A1 presented in Section (6.1), CT 502 A2 presented in Section (6.2), CT 502 A3 presented in Section (6.3), CT 502 B1 presented in Section (6.4), CT 502 B2 presented in Section (6.5), CT 502 B3 presented in Section (6.6), every segment of interest, i.e. the LD segment and the three segments corresponding to the DD, before, during and after treatment, the PC amplitude vector is obtained using the proposed method and compared with the FFT estimation, the standard method used today in chronobiology.

## 6.1 CT 502 A1: Photon Absorption and Activity

In the following we present the results of the proposed method compared with the standard method used today for analysing the chronobiological signals, the FFT method, on real data, obtained in experiments performed in male KI/KI Per2::luc mice, aged 10 weeks. This section is dedicated to the complete study, i.e the study of the physiology, corresponding to the segment LD (light-dark, 12h-12h), where the mouse was synchronized and DD (total darkness), the segment corresponding to the **before** treatment and the study of the behaviour **during** and **after** treatment. For the study, two signals were recorded: the **Activity Signal** measuring the activity (movement) of the mouse and the **Photon Absorption Signal**, measuring the number of photons absorbed.

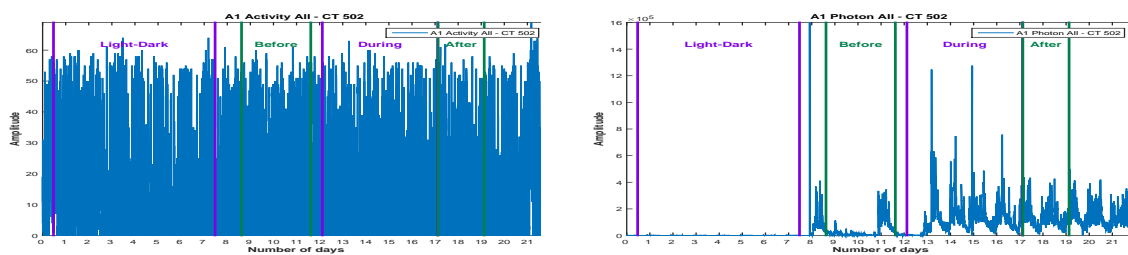


Figure 6.1: CT 502, A1 Mouse: Activity (a) and Photon Absorption (b) raw data

Figure (6.1) presents the raw data corresponding to the Activity signal and Photon Absorption signal. The signals were recorded for 22 days, sampled every minute. For each signal, we consider 4 segments: **LD Period**, corresponding to the synchronization part, then three parts corresponding to the **DD period**: before, during and after treatment. We are interested in the periodic component (PC) vector corresponding to each signal and in the corresponding acrophase. The stability of the period is verified using the moving window strategy, where the available data allows.

### 6.1.1 Dominant Period evolution: Activity

We start the analysis with the signal corresponding to the Activity. We consider zero-mean signals, normalized between [-10:10] and sampled every hour.

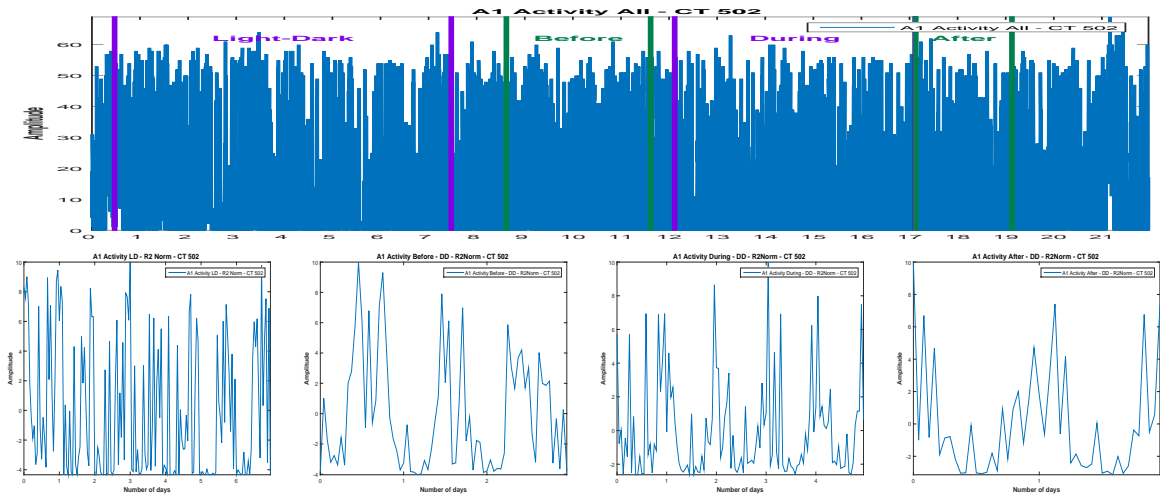


Figure 6.2: Activity raw data (a) and the corresponding parts, (b), (c), (d), normalized and one-hour sampled

For the LD part, 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 principal period

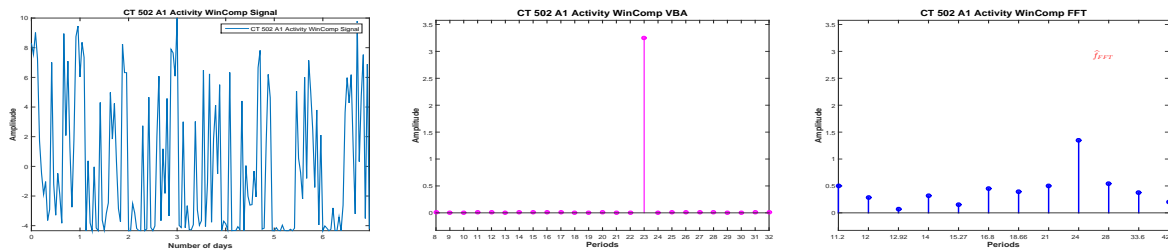


Figure 6.3: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

is estimated at 24 hours. Evidently, beside the incertitude associated with the FFT estimated PC, the existence of other rhythms can't be established, being difficult to interpret all the picks present in the estimated PC. Via the Proposed Method, the principal period is estimated at 23 hours. The amplitude corresponding to the 23h period via the Proposed Method is much higher.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method. Via the Proposed Method, three windows present a 23 hours periodicity and for one window, a 24 hours periodicity is detected. Via FFT method the principal period seems stable, every window presenting a dominant period of 24 hours. This result is consistent with the PC estimation via Proposed method, when all the signal was considered.

The stability or variability of the PC amplitudes vector for the LD part of the Activity signal is presented in Figure (6.5):

## 6.1.1 - Dominant Period evolution: Activity

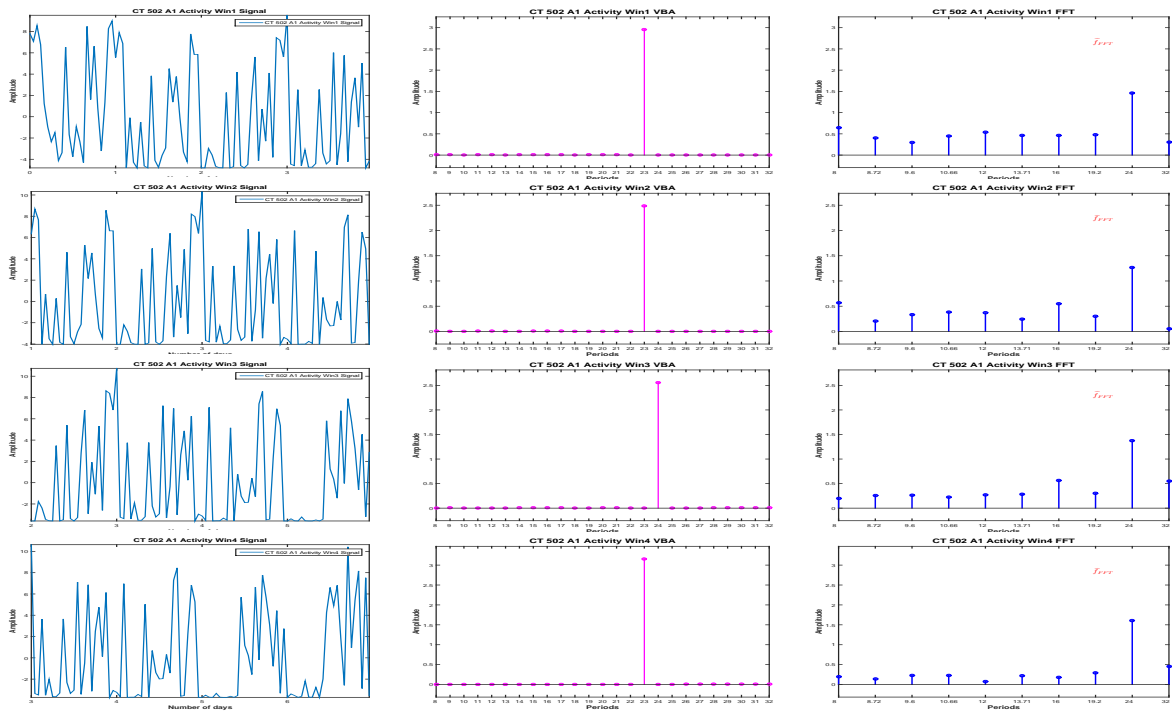


Figure 6.4: PC Stability: PC estimation via FFT and VBA for 4-days length signals

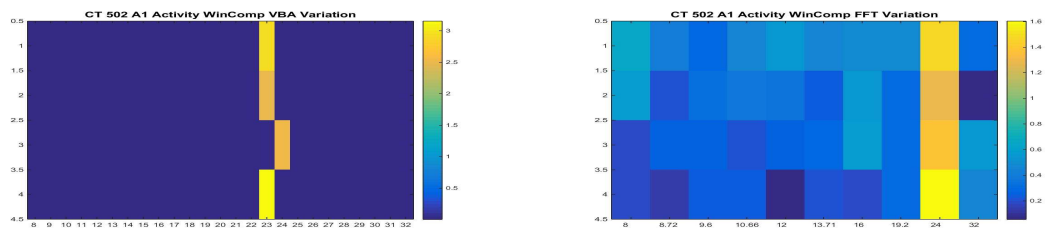


Figure 6.5: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in the Figure (6.6):

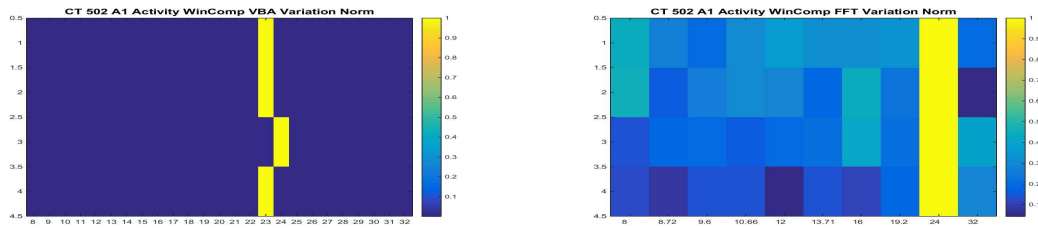


Figure 6.6: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

We consider now the DD period. First, we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, (6.7): Via the FFT, the highest pick is set at 24 hours and the next three highest

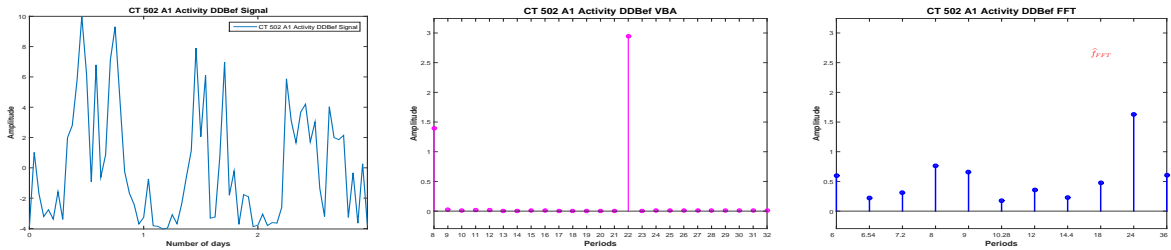


Figure 6.7: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

picks are set at 8, 36 and 6 hours. Of course, 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 result presents a very important uncertainty. This means that for this particular case, even the classification of the rhythm (circadian or infradian) is uncertain via FFT. Via the Proposed method the dominant period is set at 22 hours and the other pick present in the PC vector correspond to 8 hours.

We consider the **during** treatment part of the data, for which five days are available, Figure (6.8). Via the proposed method, the estimated PC vector is a sparse vector, in accordance

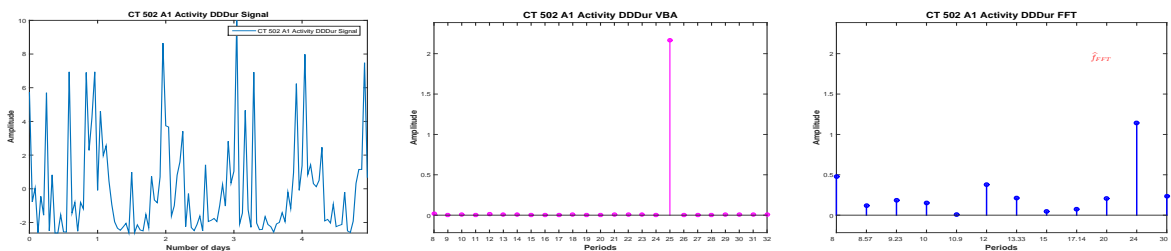


Figure 6.8: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

with the model and the dominant period is estimated at 25 hours. For the FFT estimated PC vector, the dominant period is estimated at 24 hours, but the estimation is not sparse. In this case, the next two important picks correspond to 8 and 12 hours, harmonics of 24 hours. We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method. In this case, analysing the **during**

### 6.1.1 - Dominant Period evolution: Activity

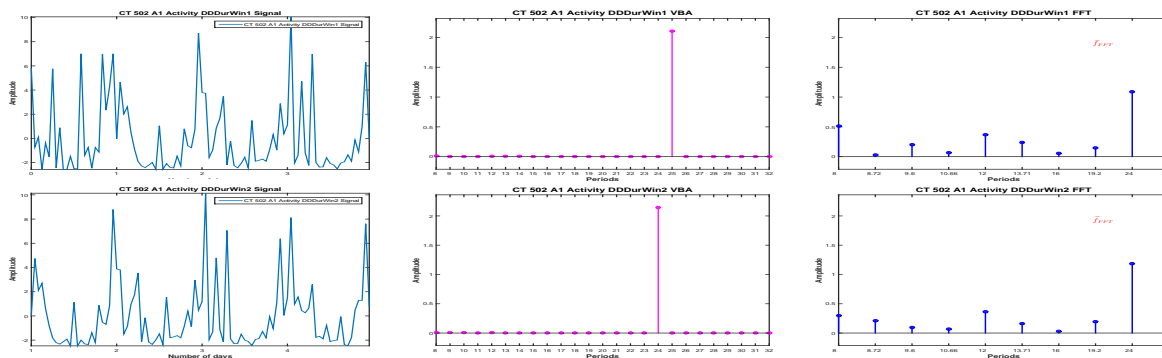


Figure 6.9: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during

segment using the moving window we note that via the Proposed Method a variability of the dominant period is detected, while via FFT the dominant period seems stable. This can be viewed in Figure (6.10). An image of dominant period stability or variability, regardless the

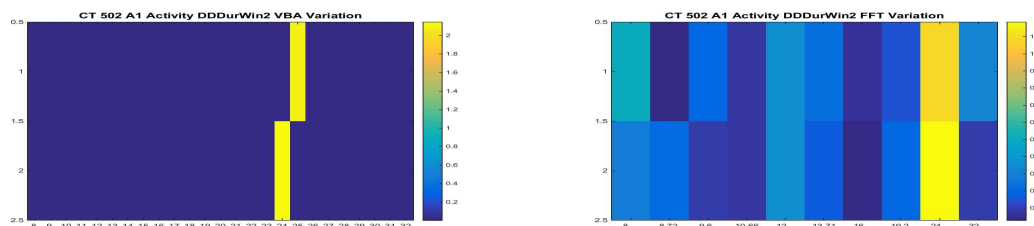


Figure 6.10: PC Stability: Proposed Method (a) vs. FFT (b)

numerical value of the amplitude associated with the dominant period for **during** segment is presented in Figure (6.11):

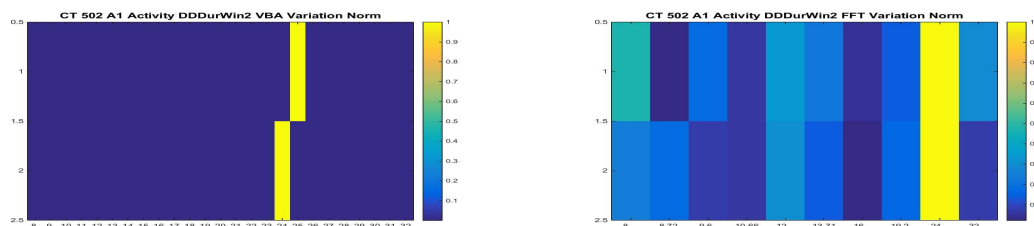


Figure 6.11: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

For the **after** treatment part of the data 2 days are available. Via the FFT estimation, the dominant period is 24 hours, while via the proposed method the dominant period is 25 hours.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Activity signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.1).

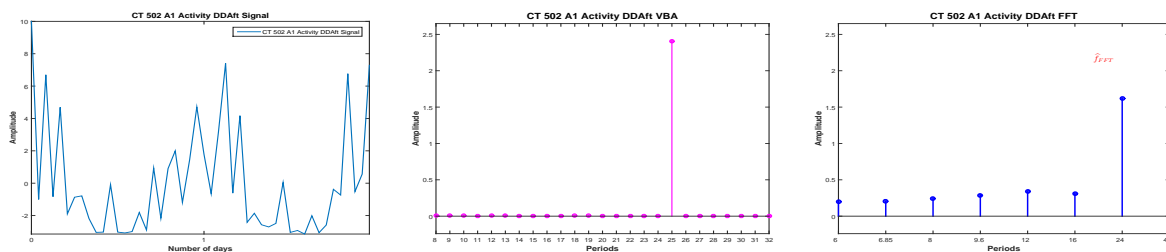


Figure 6.12: DD after signal (a) and the corresponding PC via VBA (b) and FFT (c)

CT 502 - A1	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	23	24
LD win 1	23	24
LD win 2	23	24
LD win 3	24	24
LD win 4	23	24
DD before	22	24
DD during	25	24
DD during win 1	25	24
DD during win 2	24	24
DD after	25	24

Table 6.1: CT 502 A1: Activity Dominant Period Stability: Proposed Method vs. FFT

### 6.1.2 Dominant Period evolution: Photon

In this section we analyse the signal corresponding to the photon absorption, i.e. the gene expression. Like before, we consider mean-zero signals, normalized between  $[-10;10]$  and sampled every hour. In this case, the analysis corresponding to the LD part has no biological interest, since the RT-BIO photomultiplier will not record the mouse liver bioluminescence but the light inside cage corresponding to the 12 hours of light.

The complete Photon absorption signal and the four segments of interest are presented in Figure (6.13).

For the LD part, 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 principal period is estimated at 24 hours. In particular, in this case, the FFT method estimates a sparse PC amplitudes vector, and along with the 24 hours, a non-zero peak is detected corresponding to 8 hours. Very similar results are obtained in this case via the Proposed Method: the principal period is also estimated at 24 hours, together with the harmonic of 8 hours. However, the amplitude corresponding to the Proposed Method is much higher.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.15). For all the windows, a 24 hours periodicity is confirmed, via the two methods, with a harmonic corresponding to 8 hours, Figure (6.16). Once again, for this segment, there is no biological interpretation.



## 6.1.2 - Dominant Period evolution: Photon

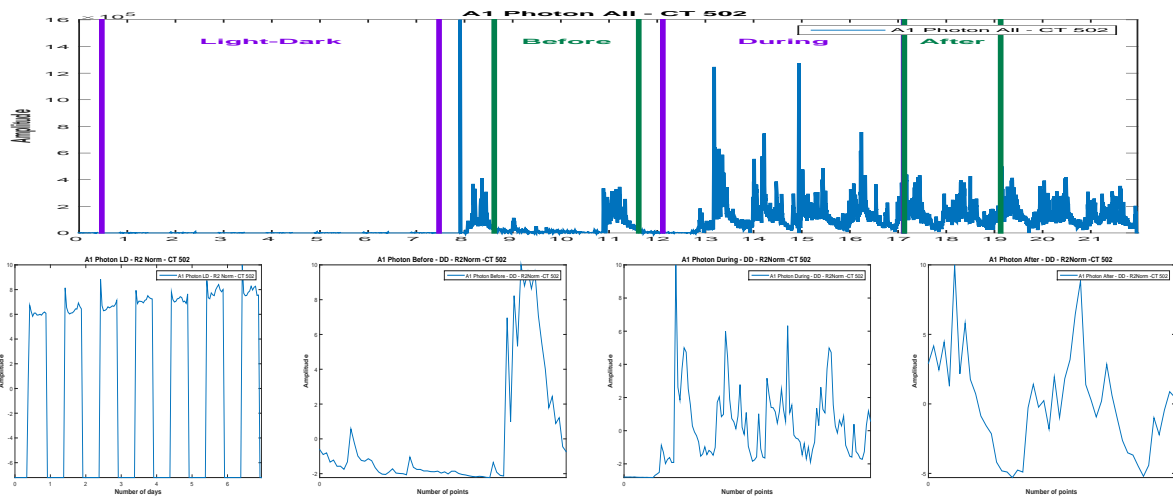


Figure 6.13: Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled

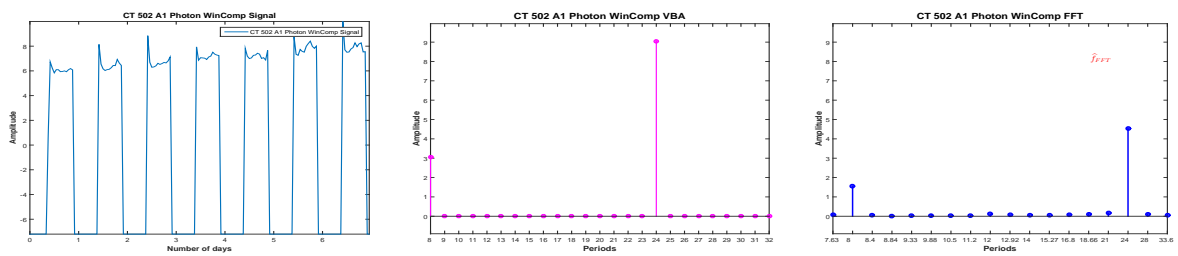


Figure 6.14: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

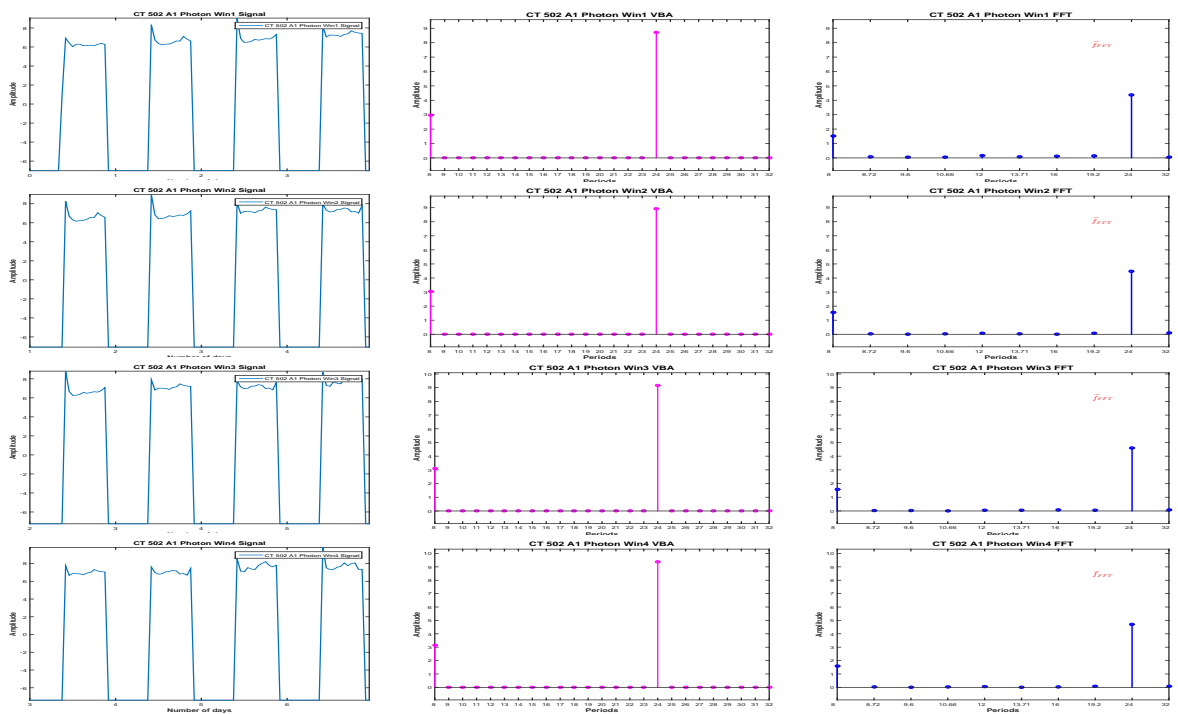


Figure 6.15: PC Stability: PC estimation via FFT and VBA for 4-days length signals

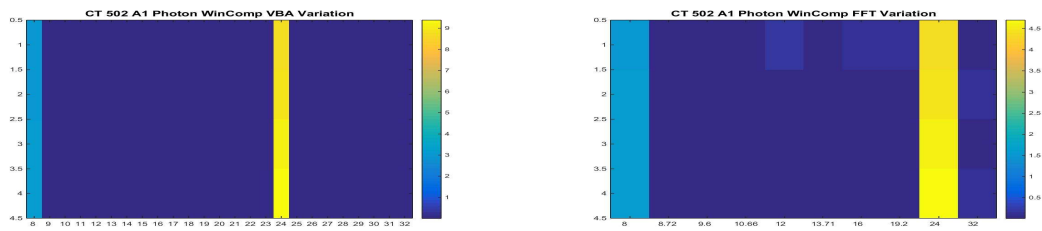


Figure 6.16: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.17):

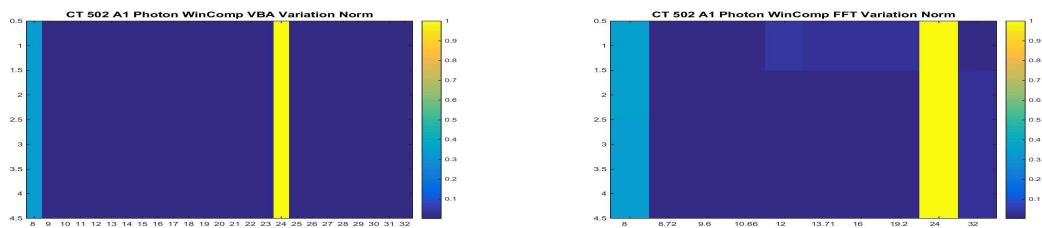


Figure 6.17: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

We consider now the DD period. First, we consider the **before** treatment segment. However, in this case the signal is corrupted, Figure (6.18):

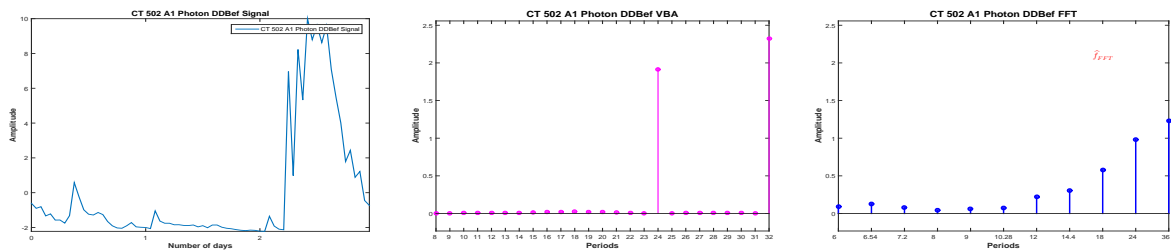


Figure 6.18: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

We consider the **during** treatment part of the data. Five days are available, Figure (6.19). Via FFT method, the dominant period is estimated at 24 hours, with important peaks corresponding to 17, 10 and 30 hours. Via the Proposed Method, the PC amplitudes vector detects three peaks: the dominant period is set at 26 hours and non-zero peaks are also detected for 17 and 23 hours.

The analysis of the **during** segment using the moving window is presented in Figure (6.20).

The visual representation of the variability of the PC amplitude vector via the two methods is presented in Figure (6.21).

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.22).

We consider the **after** treatment part of the data, two days length signal, Figure (6.23). In

## 6.1.2 - Dominant Period evolution: Photon

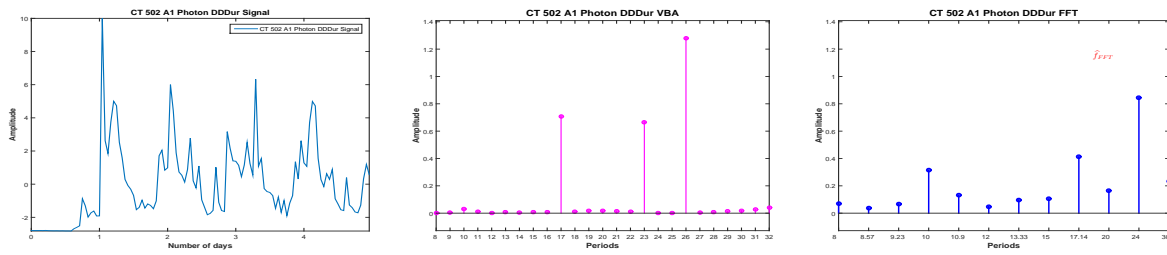


Figure 6.19: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

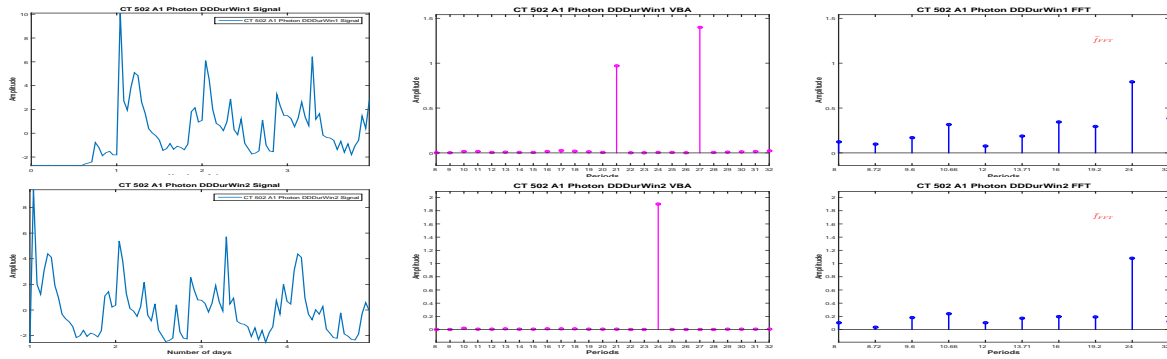


Figure 6.20: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during

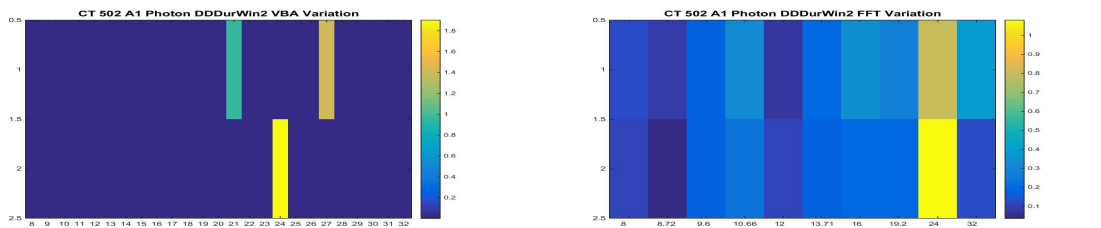


Figure 6.21: PC Stability: Proposed Method (a) vs. FFT (b)

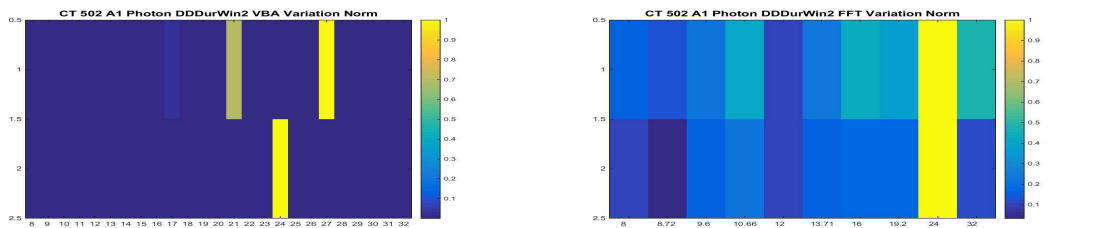


Figure 6.22: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

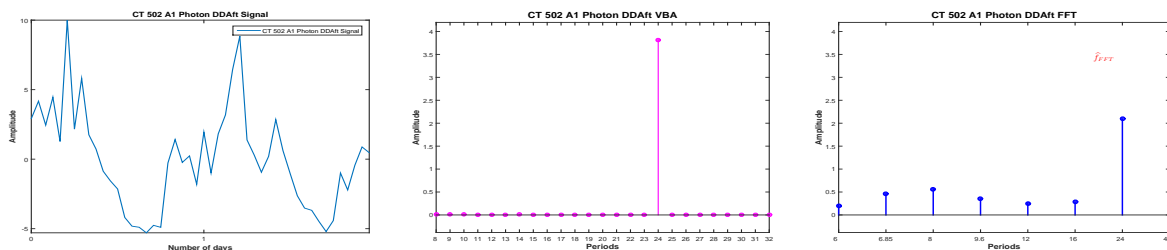


Figure 6.23: DD **after** signal (a) and the corresponding PC via VBA (b) and FFT (c)

CT 502 - A1	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	24	24
LD win 1	24	24
LD win 2	24	24
LD win 3	24	24
LD win 4	24	24
DD before	-	-
DD during	-	-
DD during win 1	-	-
DD during win 2	24	24
DD after	24	24

Table 6.2: CT 502 A1: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT

this case, the dominant period is estimated at 24 hours both via FFT and the Proposed Method.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Photon absorption signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.2).

### 6.1.3 Conclusions: Dominant Period evolution

We have compared the FFT and Proposed Method PC amplitude vector evolution during the four segments of interest for mouse CT 503 A1, considering the gene expression (Photon Absorption signal) and the rest-activity patters (Activity signal), Figure (6.1). For both data, we have considered four segments of interest, i.e. the LD part of the signal, corresponding to the synchronization and the DD part of the signal, corresponding to the complete darkness, for which we have considered the segments corresponding to the **before**, **during** and **after** treatment phase, Figure (6.2) (rest-activity patterns) and Figure (6.13).

In particular, for the first segment of interest, i.e. the LD part of data and for the **during** treatment part, we have considered four days length signals to analyse the stability or variability of the dominant period for the considered segment itself.

For the gene expression (photon absorption signal), we have presented the results corresponding to each segment, but we haven't considered the DD **before** and **during** treatment parts of the data, since the corresponding signals are corrupted. The dominant period evolution, corresponding to the Proposed Method and FFT, is presented in Table (6.2). In this case, via both methods, the dominant period is 24 hours and is stable during the signal. We note that while the FFT approach do not give a sufficient precision, the results are in accordance, in the sense that generally, the peaks that appear in the PC amplitude vector estimated via the Proposed Method, have corresponding important peaks in the FFT PC amplitude vector. For the gene expression data, for the LD part, we note that both methods detect the dominant period at 24 hours and the 8 hours harmonic.

The dominant period evolution, corresponding to the Proposed Method and FFT for the

### 6.2.1 - Dominant Period evolution: Activity

---

rest-activity patterns (Activity) signal is presented in Table (6.1). In this case, via FFT method the dominant period is 24 hours for each segment and the evolution of the dominant period seems to be stable. However, via the Proposed Method, the dominant period has certain variation around 24 hours.

In particular, via the Proposed Method, for the rest-activity patterns, the dominant period for the LD segment was estimated at 23 hours, Figure (6.3) (for three windows corresponding to LD part of the signal, the dominant period is estimated at 23 hours and for one window is estimated at 24 hours, Figure (6.5)). Furthermore, for the first three days of complete darkness, corresponding to the **before** treatment part of the signal, via the Proposed Method the circadian rhythm seems to be perturbed: the dominant period is estimated at 22 hours, Figure (6.7) and then, for the **during** and **after** treatment parts of the data, the dominant period becomes 25 hours, Figure (6.8) and Figure (6.12).

## 6.2 CT 502 A2: Photon Absorption and Activity

---

This section presents the comparison between the Proposed Method and the FFT method, for another experiment, performed in the same conditions as the one presented in Section (6.1).

For the study, two signals are recorded: the **Activity** signal measuring the rest-patterns (movement) of the mouse and the **Photon absorption** signal, measuring the number of photons absorbed, i.e. the gene expression. Figure (6.24) presents the raw data corresponding

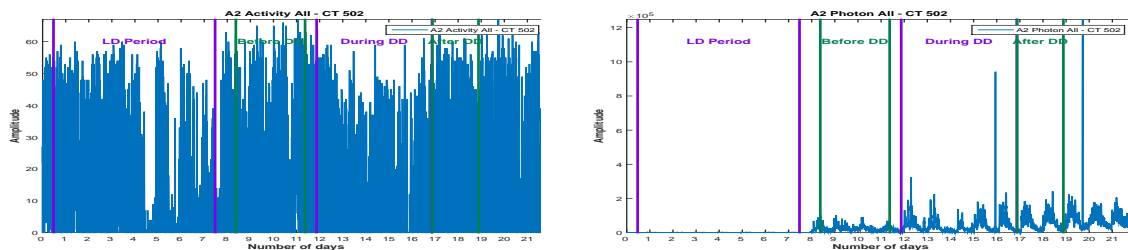


Figure 6.24: CT 502, A2 Mouse: Activity (a) and Photon Absorption (b) raw data

to the Activity signal and Photon Absorption signal. The signals were recorded for 22 days, sampled every minute. Like in Section (6.1), for each signal, we consider the same 4 segments: **LD Period**, corresponding to the synchronization part, then three parts corresponding to the **DD period**: before, during and after treatment. We are interested in the periodic component (PC) vector corresponding to each signal. The stability of the period is verified using the moving window strategy, where the available data allows.

### 6.2.1 Dominant Period evolution: Activity

We present the analysis with the signal corresponding to the Activity. We will consider mean-zero signals, normalized between [-10:10] and sampled every hour, Figure (6.25).

For the LD part, 7 days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.26): Via the FFT method, two principal periods are detected: 24 and 21 hours. Via the Proposed method, the two principal periods detected are 23 and 21 hours.

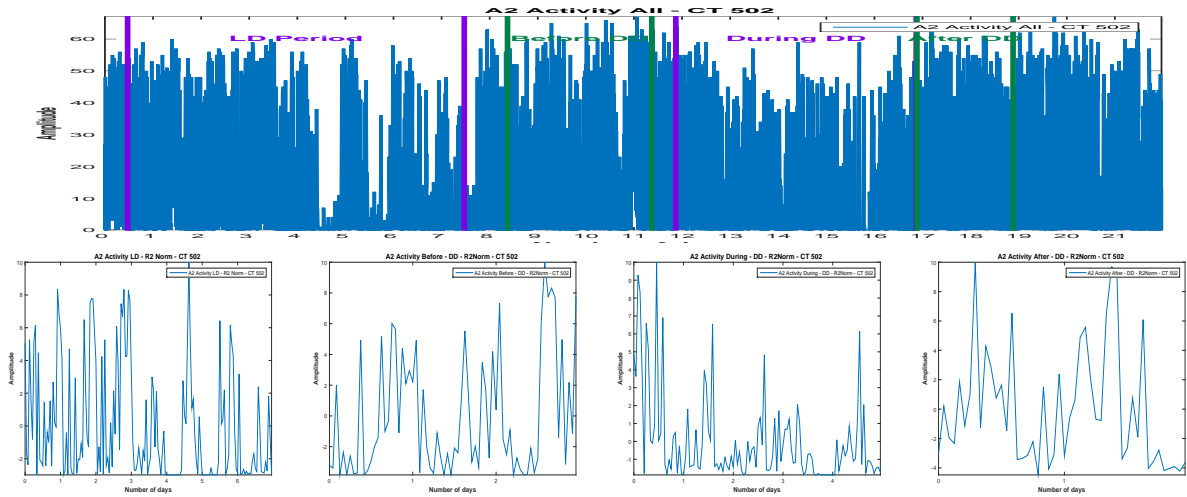


Figure 6.25: Activity raw data (a) and the corresponding parts, (b), (c), (d), normalized and one-hour sampled

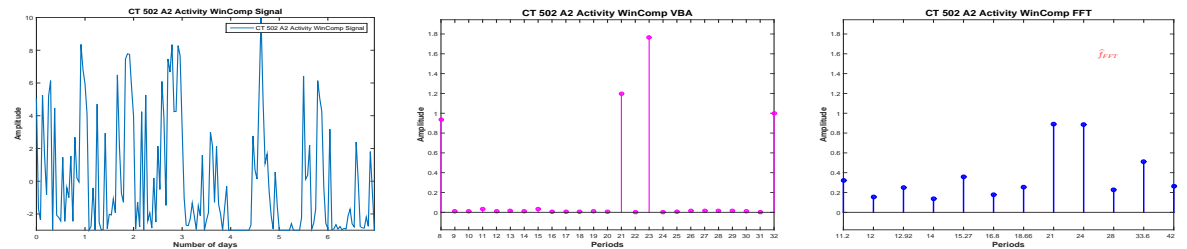


Figure 6.26: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

We consider four days length signals (windows) from the available LD signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.27). Via FFT method, the dominant period is 24 hours for all windows, therefore stable. Via the Proposed Method, the dominant period is different for each window: 21, 22, 23 and 25 hours, Figure (6.28). An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.29): For the DD period, we consider first the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, Figure (6.30): The estimation corresponding to the FFT method, the dominant peak corresponds to 24 hours, and the next important peak corresponds to 14 hours. For the estimation corresponding to the Proposed Method, the PC amplitude vector is sparse, with two non-zero peaks, corresponding to 24 and 14 hours.

For the **during** treatment part of the data, five days are available. The comparison between the Proposed Method and FFT is presented in Figure (6.31). The estimation corresponding to the FFT method sets the dominant period at 12 hours, with important peaks corresponding to 30, 24 and 12 hours. Via the Proposed Method, the dominant period is set at 32 hours, and the other non-zero amplitudes in the PC amplitudes vector corresponds to 25 and 12 hours. We note that in this case, a certain consistency between the two results is shown: both methods detect an important peak at 12 hours, an important peak around 24 hours (25 hours for the Proposed Method) and the dominant period in the infradian domain (i.e. superior to 30 hours).

The analysis of the **during** segment using the moving window is presented in Fig-

## 6.2.1 - Dominant Period evolution: Activity

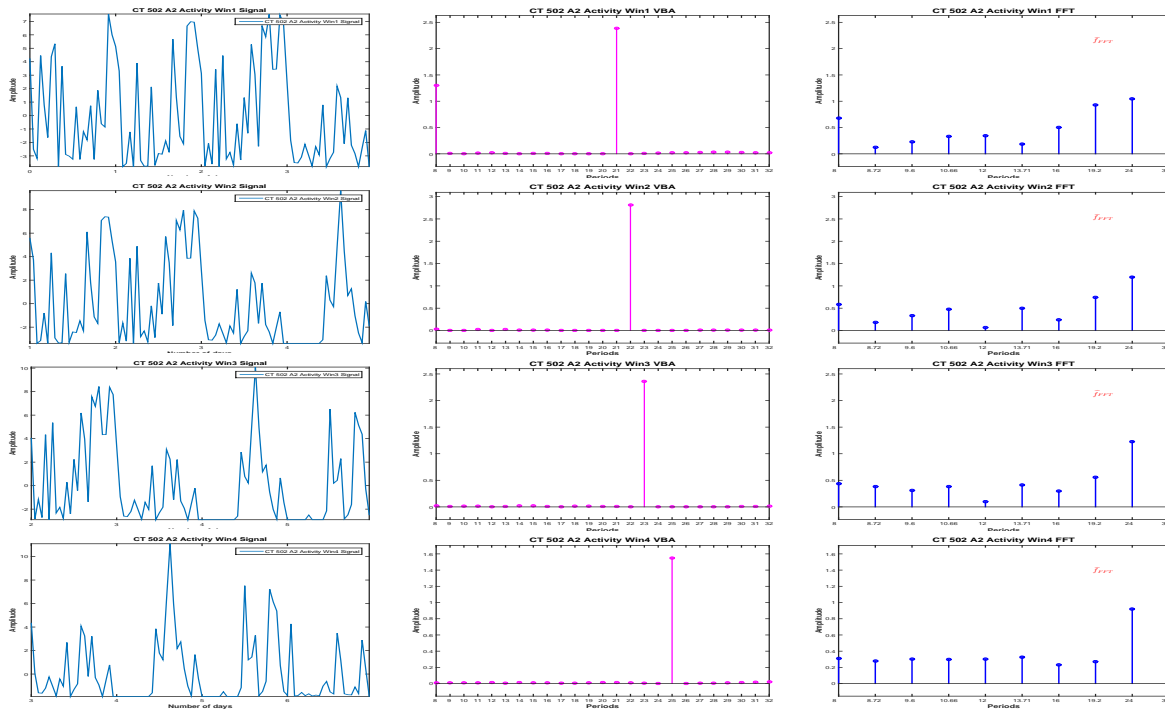


Figure 6.27: PC Stability: PC estimation via FFT and VBA for 4-days length signals

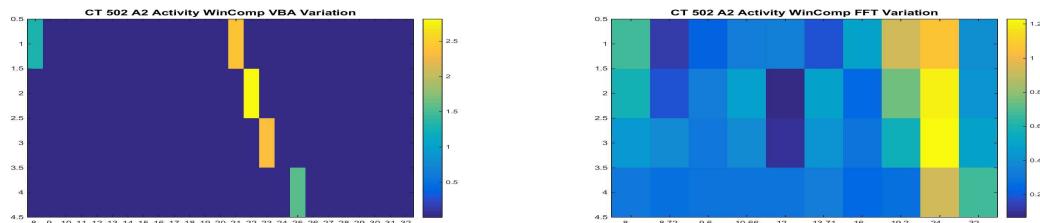


Figure 6.28: PC Stability: Proposed Method (a) vs. FFT (b)

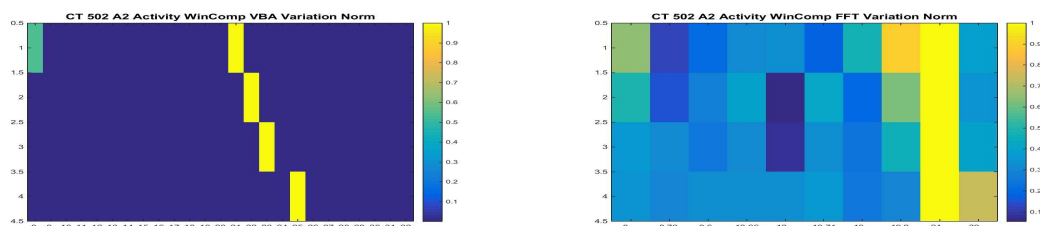


Figure 6.29: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

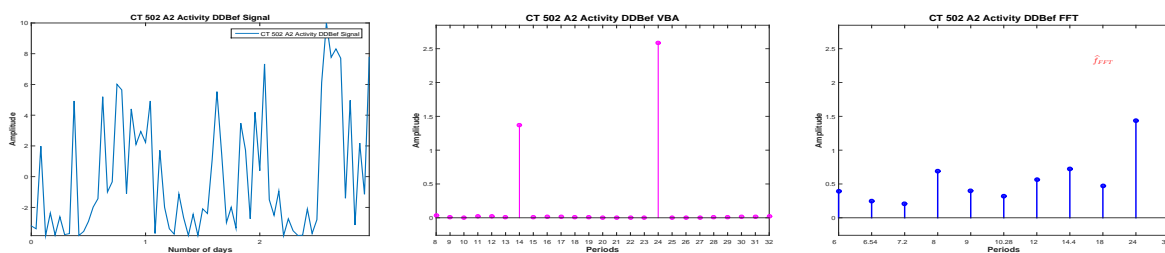


Figure 6.30: DD before signal (a) and the corresponding PC via VBA (b) and FFT (c)

## CT 502 A2: Photon Absorption and Activity

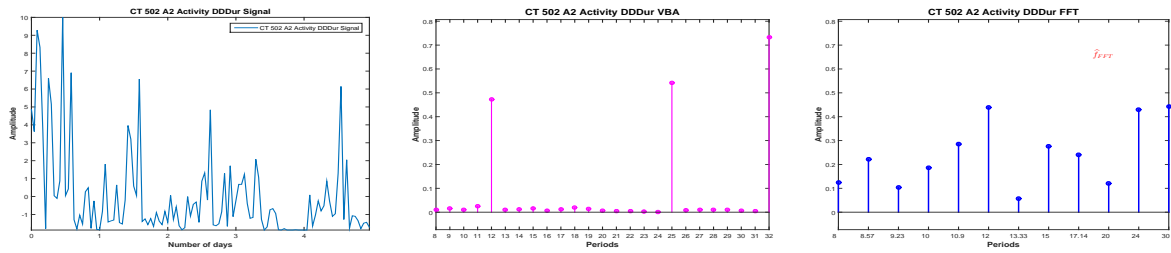


Figure 6.31: DD during signal (a) and the corresponding PC via VBA (b) and FFT (c)

ure (6.32). For the first window, via the two methods, the dominant period is 32 hours, placing

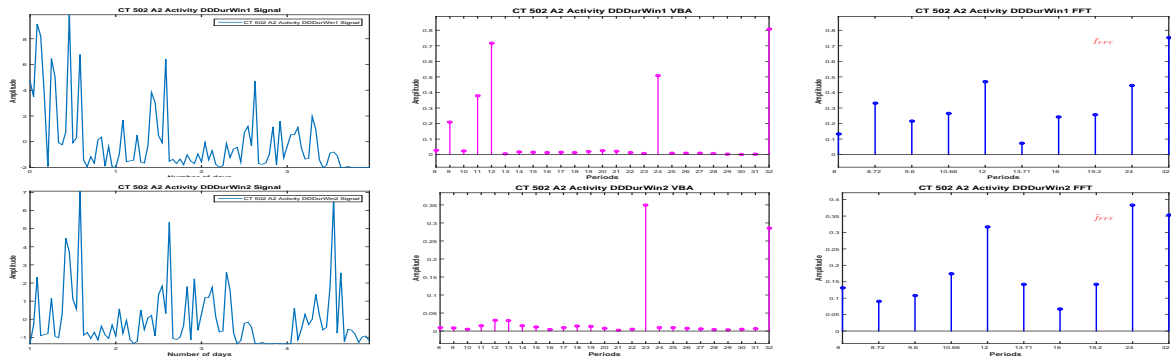


Figure 6.32: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during

the rhythm in the ultradian domain. For the second window, both methods are placing the rhythm in the circadian domain: via the FFT, the dominant period is detected at 24 hours, and via the Proposed Method at 23 hours, Figure (6.33).

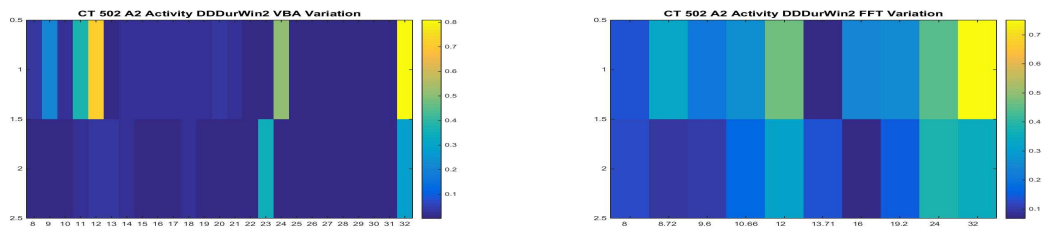


Figure 6.33: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.34):

We consider the **after** treatment part of the data, for which two days are available, Figure (6.35).

Via the Proposed Method, the dominant period is estimated at 23 hours, while via the FFT, the dominant period is estimated at 24 hours.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Activity signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.3).



## 6.2.1 - Dominant Period evolution: Activity

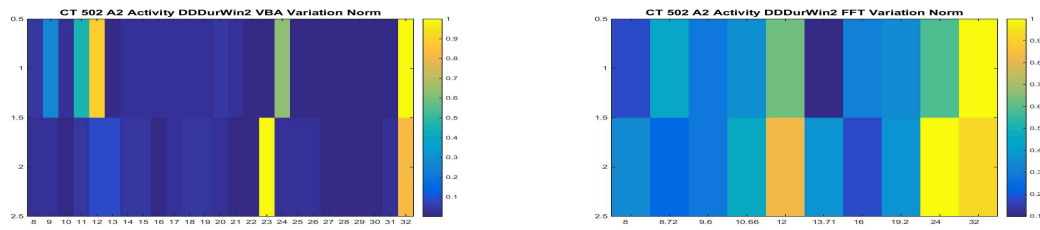


Figure 6.34: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

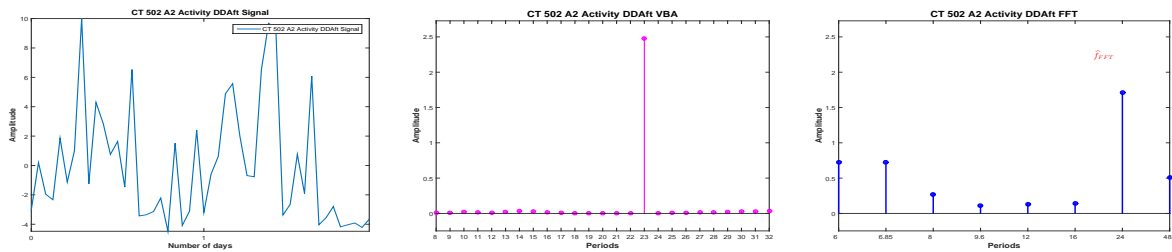


Figure 6.35: DD after signal (a) and the corresponding PC via VBA (b) and FFT (c)

CT 502 - A2	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	23	21
LD win 1	21	24
LD win 2	22	24
LD win 3	23	24
LD win 4	25	24
DD before	24	24
DD during	32	12
DD during win 1	32	32
DD during win 2	23	24
DD after	23	24

Table 6.3: CT 502 A2: Activity Dominant Period Stability: Proposed Method vs. FFT

### 6.2.2 Dominant Period evolution: Photon

In this section we analyse the signal corresponding to the photon absorption, i.e. the gene expression. Like before, we consider mean-zero signals, normalized between [-10:10] and sampled every hour. In this case, the analysis corresponding to the LD part has no biological interest, since the RT-BIO photomultiplier will not record the mouse liver bioluminescence but the light inside cage corresponding to the 12 hours of light. The complete Photon absorption signal and the four segments of interest are presented in Figure (6.36).

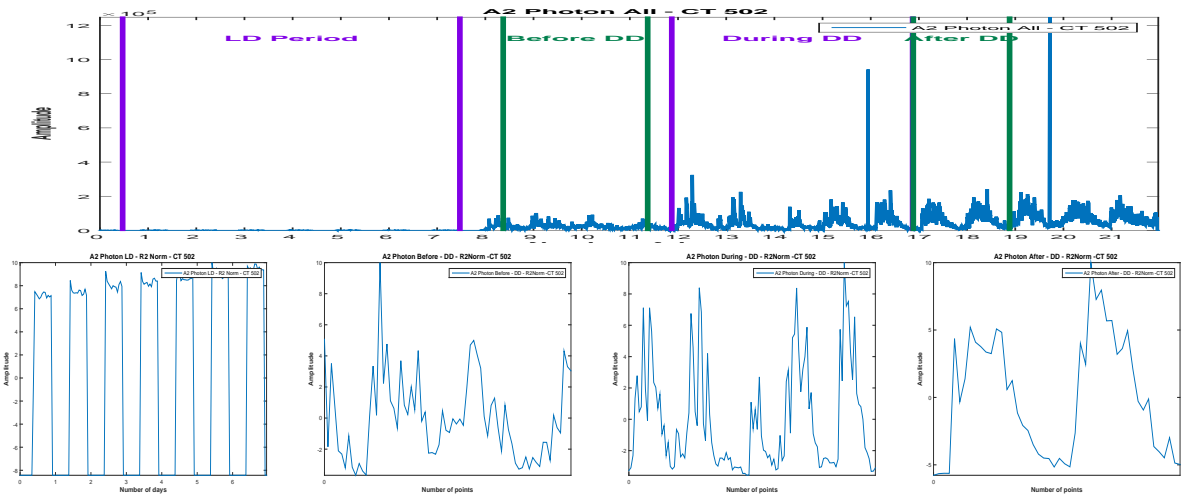


Figure 6.36: Photon raw data (a) and the corresponding parts, (b), (c), (d), normalized and one-hour sampled

For the LD part, 8 days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.37): Via the FFT method, the

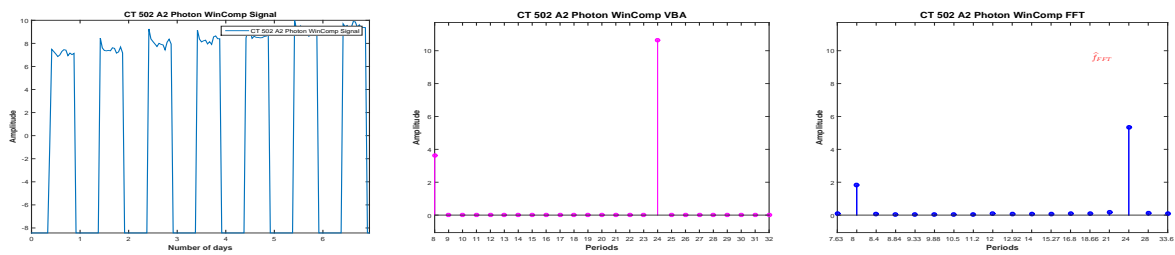


Figure 6.37: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

principal period is estimated at 24 hours. Via the Proposed Method, the principal period is also estimated at 24 hours, together with the harmonic of 8 hours. However, the amplitude corresponding to the Proposed Method is much higher.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.38).

For all the windows, a 24 hours periodicity is confirmed, via the two methods, Figure (6.39).

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.40):

## 6.2.2 - Dominant Period evolution: Photon

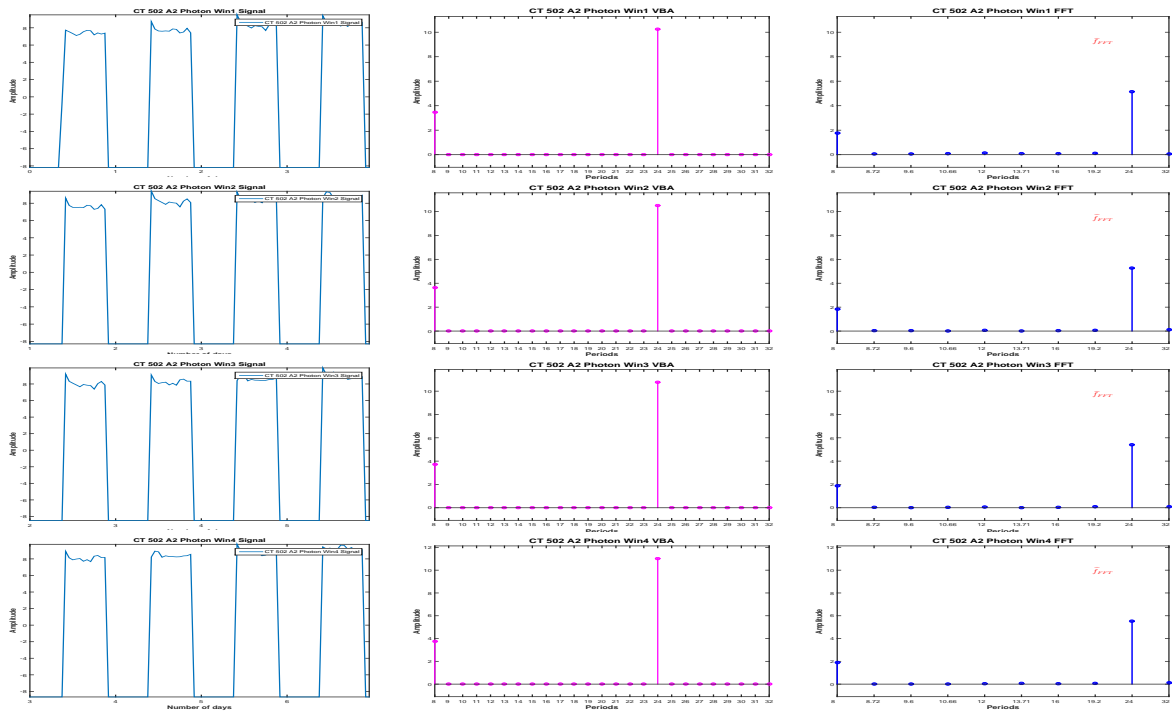


Figure 6.38: PC Stability: PC estimation via FFT and VBA for 4-days length signals

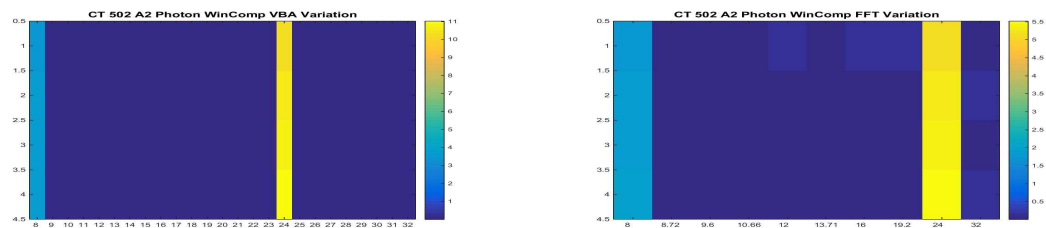


Figure 6.39: PC Stability: Proposed Method (a) vs. FFT (b)

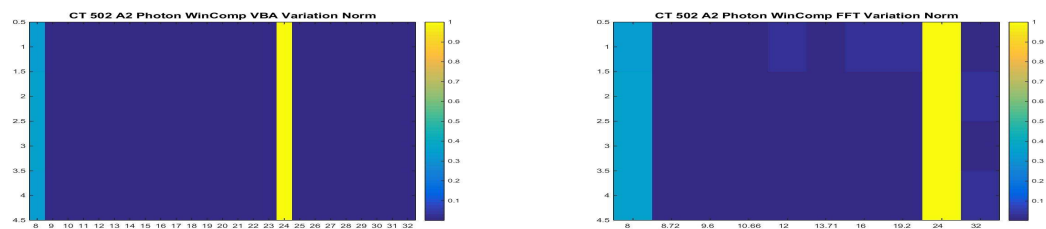


Figure 6.40: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

For the DD period, first we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, Figure (6.41):

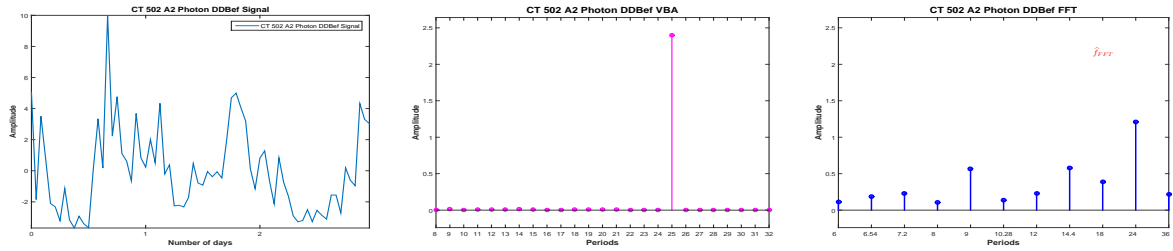


Figure 6.41: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

Via the FFT, the highest pick is set at 24. Of course, 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 25 hours.

We now consider the **during** treatment part of the data, a 5-days length signal, Figure (6.42).

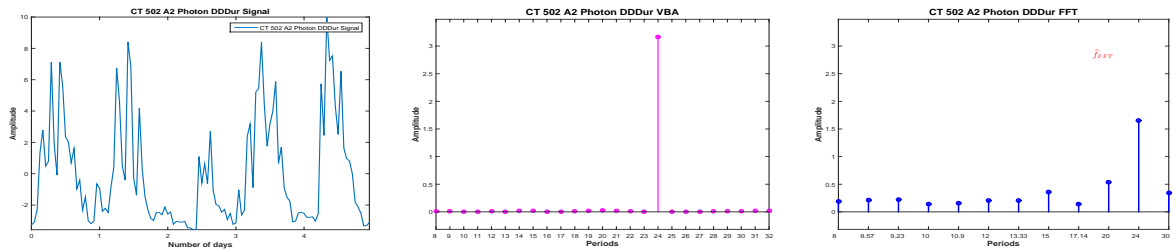


Figure 6.42: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

Both methods are confirming the 24 hours periodicity. We note again the difference between the amplitudes corresponding to the two methods and the sparse PC vector returned by the Proposed Method.

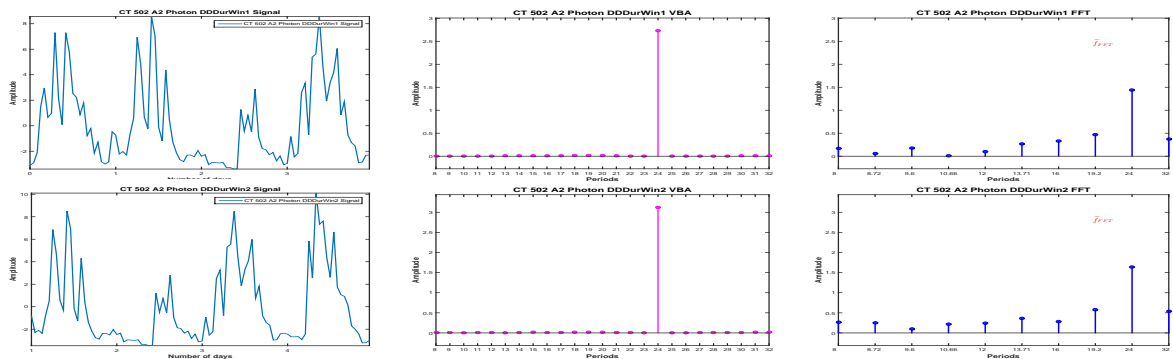


Figure 6.43: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during

Analysing the **during** segment using the moving window, the 24 hour periodicity is

### 6.2.3 - Conclusions: Dominant Period evolution

confirmed by both methods, and also the stability of this periodicity for this segment, Figure (6.44).

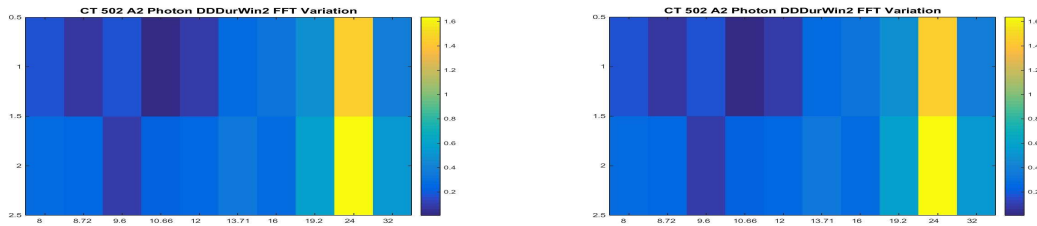


Figure 6.44: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.45):

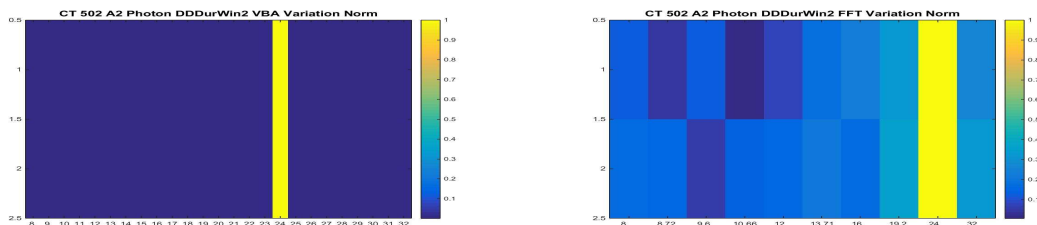


Figure 6.45: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

For the **after** treatment part of the data, a 2-days length signal is available. In this case, the

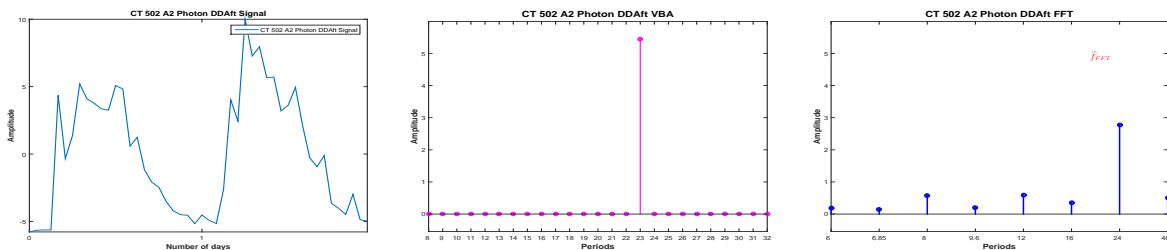


Figure 6.46: DD **after** signal (a) and the corresponding PC via VBA (b) and FFT (c)

dominant period is estimated at 24 hours via FFT. This value is uncertain, for this particular case (a 2-days length signal) the interval (16:48) having only one value that is appearing the FFT PC vector, i.e. 24 hours. Via the Proposed Method, the dominant period is set at 23 hours.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Photon absorption signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.4).

### 6.2.3 Conclusions: Dominant Period evolution

We compared the FFT and Proposed Method PC amplitude vector evolution during the four segments of interest for mouse CT 503 A2, considering the gene expression (Photon Absorption signal) and the rest-activity patters (Activity signal), Figure (6.24). For both data, we

CT 502 - A2	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	24	24
LD win 1	24	24
LD win 2	24	24
LD win 3	24	24
LD win 4	24	24
DD before	25	24
DD during	24	24
DD during win 1	24	24
DD during win 2	24	24
DD after	23	24

Table 6.4: CT 502 A2: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT

have considered four segments of interest, i.e. the LD part of the signal, corresponding to the synchronization and the DD part of the signal, corresponding to the complete darkness, for which we have considered the segments corresponding to the **before**, **during** and **after** treatment phase, Figure (6.25) (rest-activity patterns) and Figure (6.36).

In particular, for the first segment of interest, i.e. the LD part of data and for the **during** treatment part, we have considered four days length signals to analyse the stability or variability of the dominant period for the considered segment itself.

The dominant period evolution, corresponding to the Proposed Method and FFT for the gene expression data (photon absorption signal), is presented in Table (6.4). For this experiment, via FFT, the dominant period is 24 hours for each segment, and seems stable. Via the Proposed Method, slight variations are detected during the experiment: a 23 hour periodicity corresponding to the **before** treatment part of data and a 25 hour periodicity corresponding to the **after** treatment part of data.

The dominant period evolution, corresponding to the Proposed Method and FFT for the rest-activity patterns (Activity) signal is presented in Table (6.3). In this case, the drawbacks of the FFT method are very clear. For the LD segment of the Activity signal, seven days were available. The dominant period of this segment is 21 hours via FFT, the FFT PC amplitude vector estimating two important peaks, corresponding to 21 and 24 hours, both of them having almost equal amplitudes, with the amplitude corresponding to the 21 hours slightly higher, Figure (6.26). However, for the evolution of the dominant period for this segment, Figure (6.27), the dominant period is estimated at 24 hours for each windows, therefore stable. Via the Proposed Method, for the entire LD segment, the most important two peaks corresponds to 23 and 21 hours, and the evolution of the dominant period during this segment is 21 hours for the first window, 22 hours for the second window, 23 hours for the third window and 25 hours for the fourth window.

Also, for the **during** treatment segment, the FFT PC amplitude vector has important peaks corresponding to 12 hours, 30 hours and 24 hours. Considering the dominant period, which is 12 hours, via FFT estimation, the rhythm for the **during** treatment is placed in the ultradian

### 6.3.1 - Dominant Period evolution: Activity

---

domain, Figure (6.31). However, studying the evolution of the PC amplitude vector for this segment, via FFT, the dominant period corresponding to the first window is 32 hours, so infradian rhythm, and corresponding to the second window is 24 hours, so circadian rhythm.

Via the Proposed Method, for the **during** treatment segment, very similar peaks are estimated in the PC amplitude vector: 32 hours, 24 hours and 12 hours, but the dominant period is 32 hours, therefore the rhythm is infradian.

## 6.3 CT 502 A3: Photon Absorption and Activity

---

In the following we consider the experiment **CT 502, Mouse A3**. This section is dedicated to the complete study, i.e the study of the physiology, corresponding to the segment LD (light-dark, 12h-12h), where the mouse was synchronized and DD (total darkness), the segment corresponding to the **before** treatment and the study of the behaviour **during** and **after** treatment. For the study, two signals were recorded: the **Activity Signal** measuring the activity (movement) of the mouse and the **Photon Absorption Signal**, measuring the number of photons absorbed, Figure (6.47).

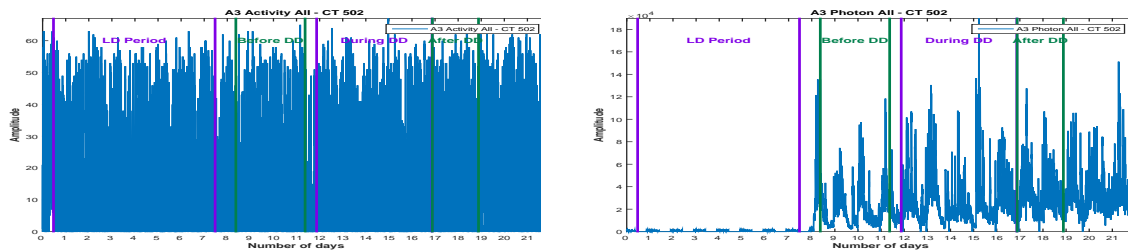


Figure 6.47: **CT 502, A3 Mouse**: Activity (a) and Photon Absorption (b) raw data

Figure (6.47) presents the raw data corresponding to the Activity signal and Photon Absorption signal. The signals were recorded for 22 days, sampled every minute. For each signal, we consider 4 segments: **LD Period**, corresponding to the synchronization part, then three parts corresponding to the **DD period**: before, during and after treatment. We are interested in the periodic component (PC) vector corresponding to each signal and in the corresponding acrophase. The stability of the period is verified using the moving window strategy, where the available data allows.

### 6.3.1 Dominant Period evolution: Activity

We start the analysis with the signal corresponding to the Activity. We will consider mean-zero signals, normalized between [-10:10] and sampled every hour. The Activity signal and the four segments of interest are presented in Figure (6.48).

For the LD part, we have 8 days available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.49). Via the FFT method, the principal period is estimated at 24 hours. The next highest pick in the PC vector is corresponding to 21 hours. Via the Proposed Method, the principal period is estimated at 22 hours, and the next highest pick in the PC vector corresponds to 20 hours

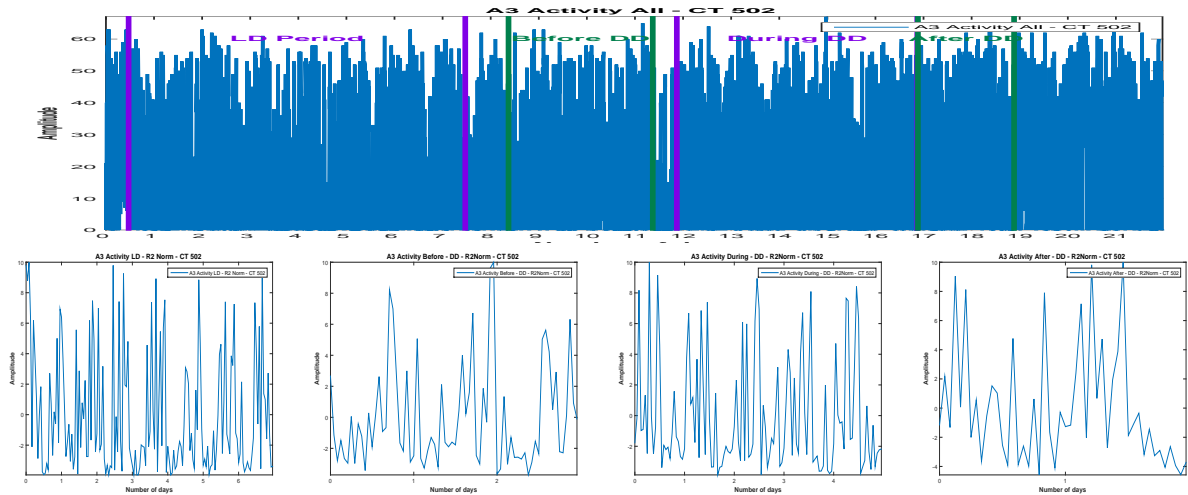


Figure 6.48: Activity raw data (a) and the corresponding parts, (b), (c), (d), normalized and one-hour sampled

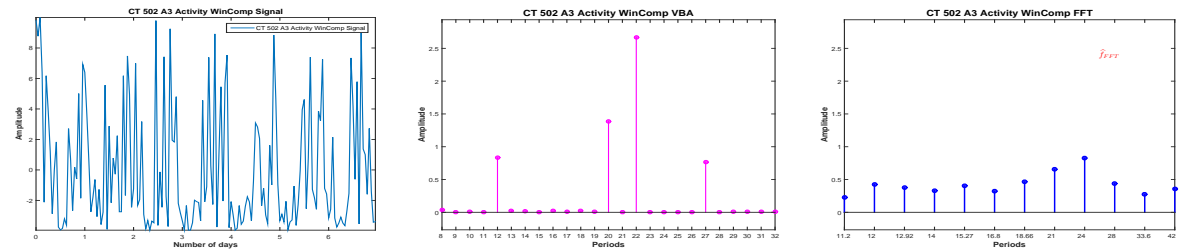


Figure 6.49: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.50).

Via FFT, the dominant period seems stable, for each window the dominant period being estimated at 24 hours. Via the Proposed Method, the dominant period shifts from 22 hours, for the first window to 23 hours for the second window to 24 hours for the last two windows, Figure (6.51).

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented Figure (6.52):.

We consider now the DD period. First, we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, Figure (6.53). Via the FFT, the highest pick is set at 24 hours and the next two highest picks are set at 7.2 and 8 hours. Of course, 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 PC amplitude vector is sparse, with only one non-zero amplitude, corresponding to 24 hours.

For the **during** treatment part of the data, 5 days are available. We compute the PC corresponding to the **during** treatment part of the data using the Proposed Method and also using the FFT, Figure (6.54). Via the Proposed Method, the estimated PC amplitude vector is a sparse vector and the dominant period is estimated at 24 hours. The other non-zero amplitudes in the PC amplitudes vector corresponds to 8, 25 and 12 hours. In the FFT estimation, the dominant period corresponds to 24 hours, and the next two highest peaks in the PC amplitude



### 6.3.1 - Dominant Period evolution: Activity

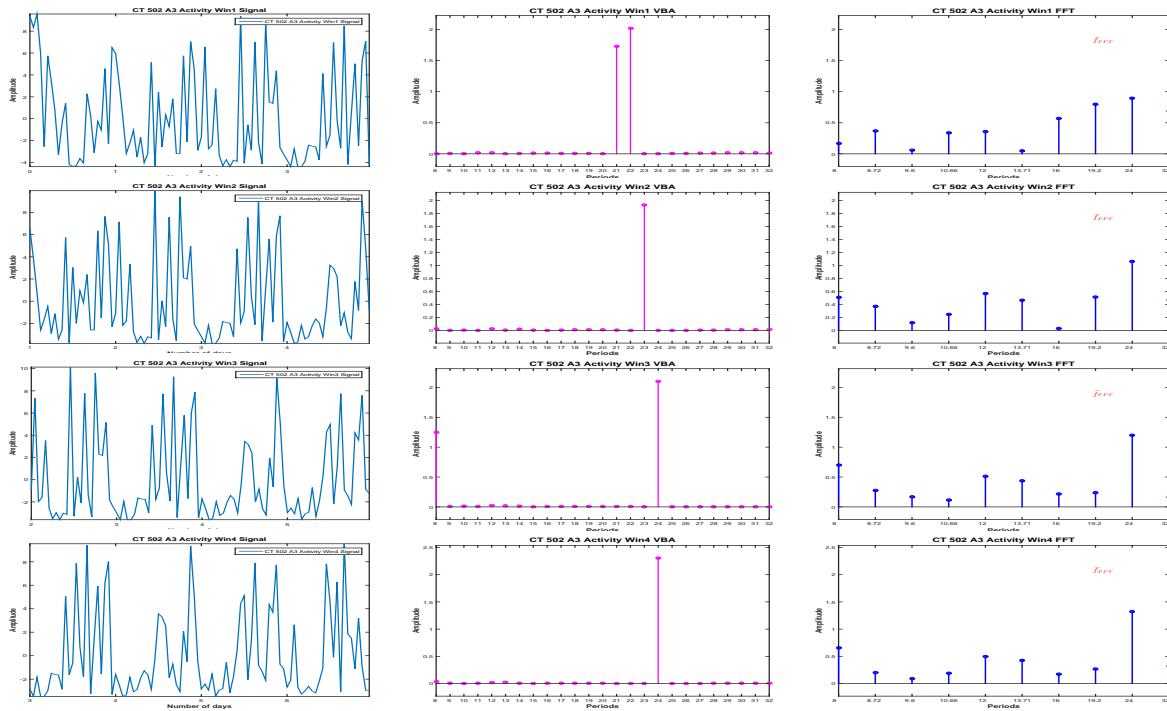


Figure 6.50: PC Stability: PC estimation via FFT and VBA for 4-days length signals

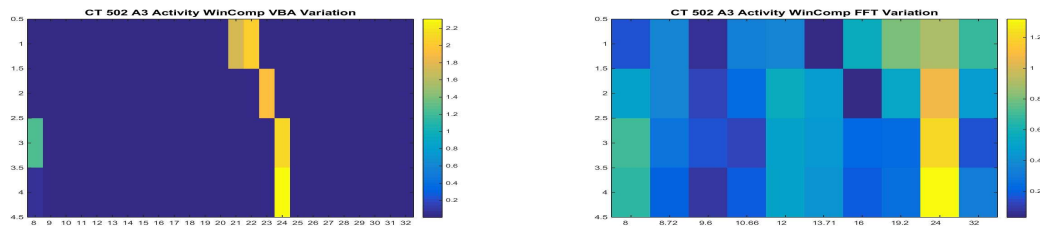


Figure 6.51: PC Stability: Proposed Method (a) vs. FFT (b)

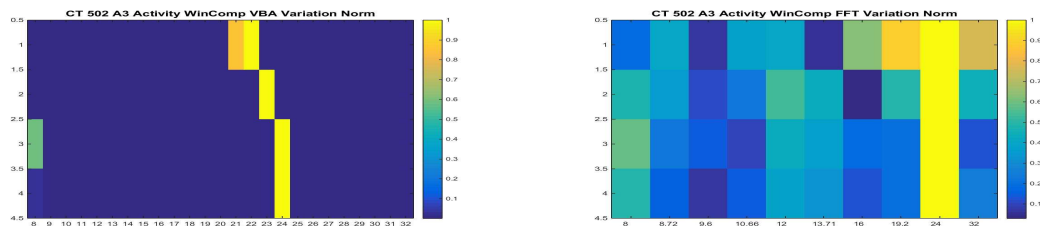


Figure 6.52: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

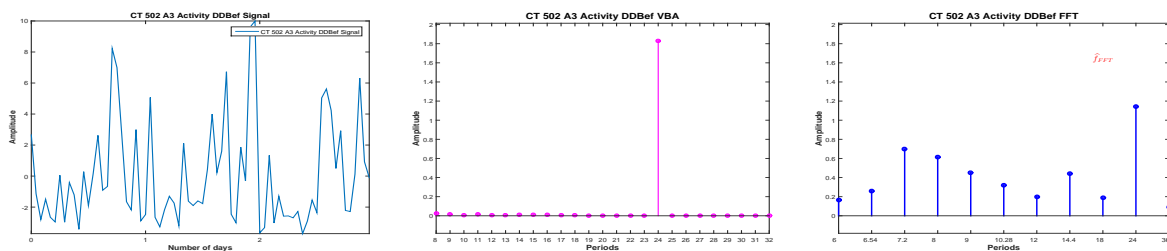


Figure 6.53: DD before signal (a) and the corresponding PC via VBA (b) and FFT (c)

## CT 502 A3: Photon Absorption and Activity

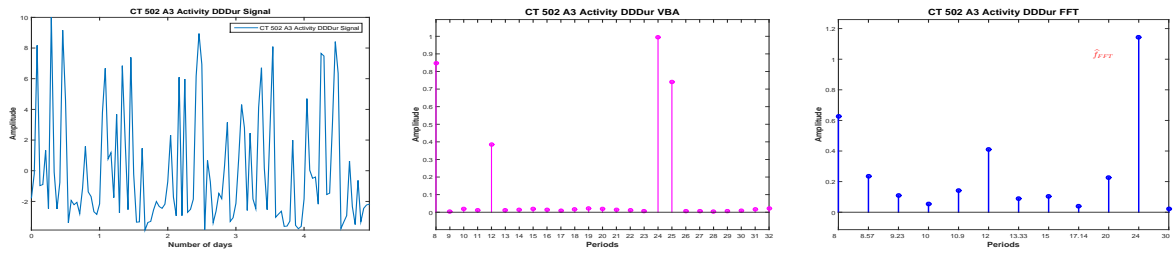


Figure 6.54: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

vector estimation corresponds to 8 and 12 hours.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.55).

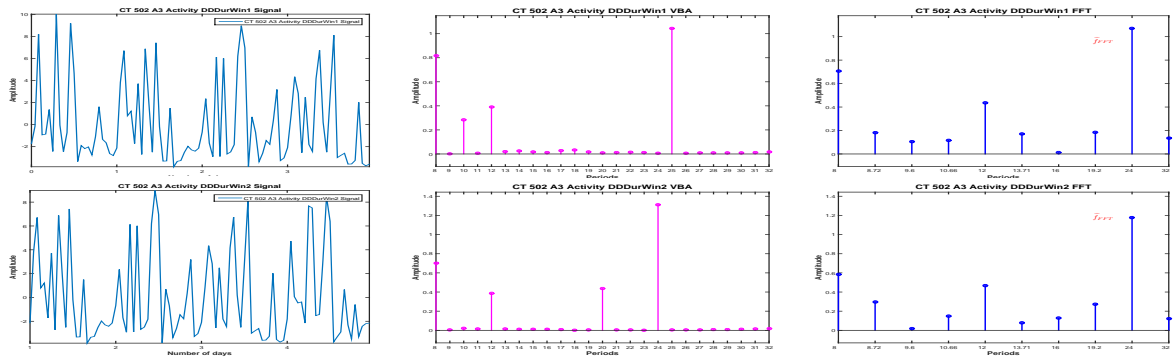


Figure 6.55: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during

We note that via the FFT, both windows are estimated with a dominant period of 24 hours, with harmonics corresponding to 8 and 12 hours. Via the Proposed Method, we find the dominant period corresponding to the first window at 25 hours (and important picks for 8 and 12 hours) and the dominant period corresponding to the second window at 25 hours (and important picks for 8 and 12 hours), Figure (6.56).

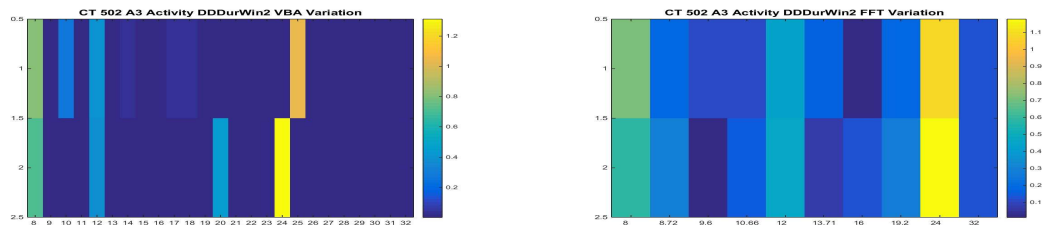


Figure 6.56: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in the following figure:

We now consider the **after** treatment part of the data. Two days are available. For the FFT, the dominant period corresponds to 24 hours. Via the Proposed Method, the dominant period is estimated at 27 hours.

### 6.3.2 - Dominant Period evolution: Photon

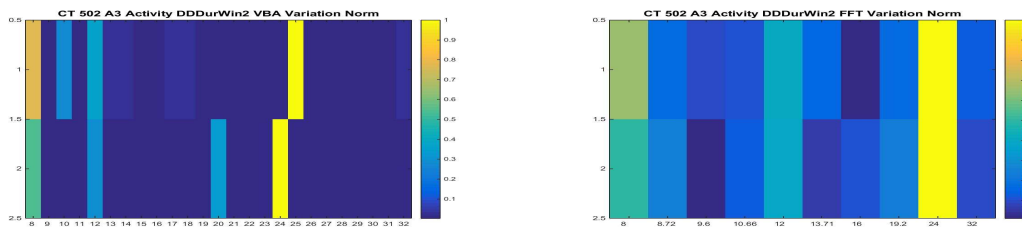


Figure 6.57: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

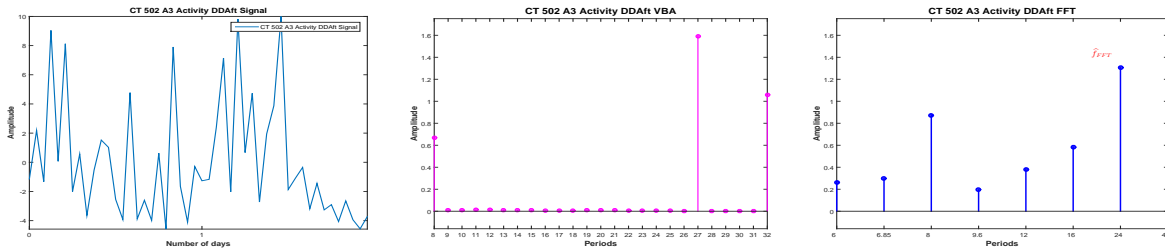


Figure 6.58: DD after signal (a) and the corresponding PC via VBA (b) and FFT (c)

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Activity signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.5).

### 6.3.2 Dominant Period evolution: Photon

In this section we analyse the signal corresponding to the photon absorption, i.e. the gene expression. Like before, we consider mean-zero signals, normalized between [-10:10] and sampled every hour. In this case, the analysis corresponding to the LD part has no biological interest, since the RT-BIO photomultiplier will not record the mouse liver bioluminescence but the light inside cage corresponding to the 12 hours of light. The complete Photon absorption signal and the four segments of interest are presented in Figure (6.59)

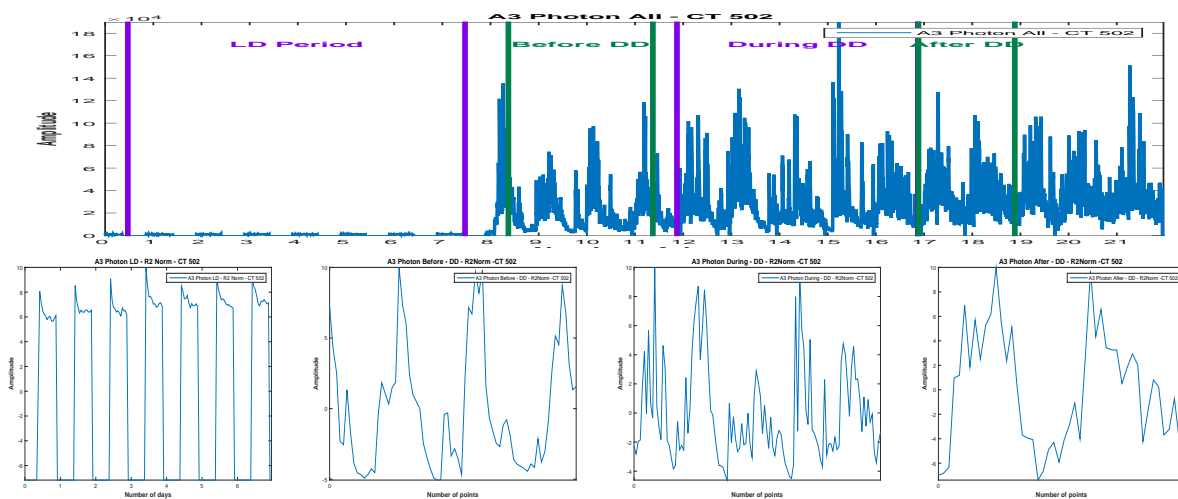


Figure 6.59: Photon raw data (a) and the corresponding parts, (b), (c), (d), normalized and one-hour sampled

CT 502 - A3	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	22	24
LD win 1	22	24
LD win 2	23	24
LD win 3	24	24
LD win 4	24	24
DD before	24	24
DD during	24	24
DD during win 1	25	24
DD during win 2	24	24
DD after	27	24

Table 6.5: CT 502 A3: Activity Dominant Period Stability: Proposed Method vs. FFT

For the LD part, 8 days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.60): Via the FFT method, the

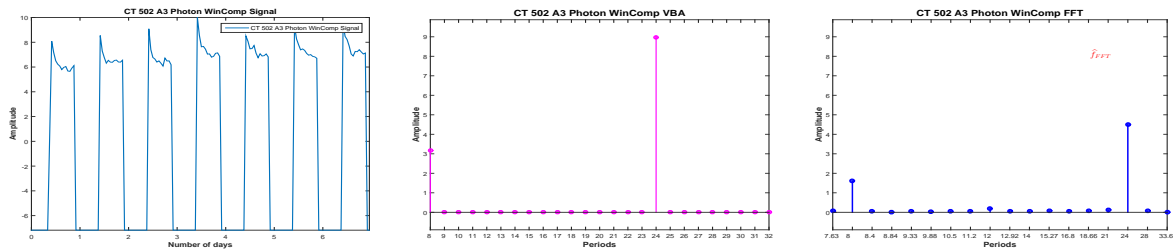


Figure 6.60: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

principal period is estimated at 24 hours and the only other non-zero peak in the PC amplitude vector corresponds to 8 hours. The dominant period of 24 hours and the corresponding 8 hours harmonic are also the only non-zero amplitudes in the PC amplitude vector corresponding to the Proposed Method.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.61).

For all the windows, a 24 hours periodicity is confirmed, via the two methods, Figure (6.62).

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.63):

We consider now the DD period. First, we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, Figure (6.64): Via the FFT, the highest pick is set at 24. Of course, 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 23 hours.

For the **during** treatment part of the data, five days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.65):

### 6.3.2 - Dominant Period evolution: Photon

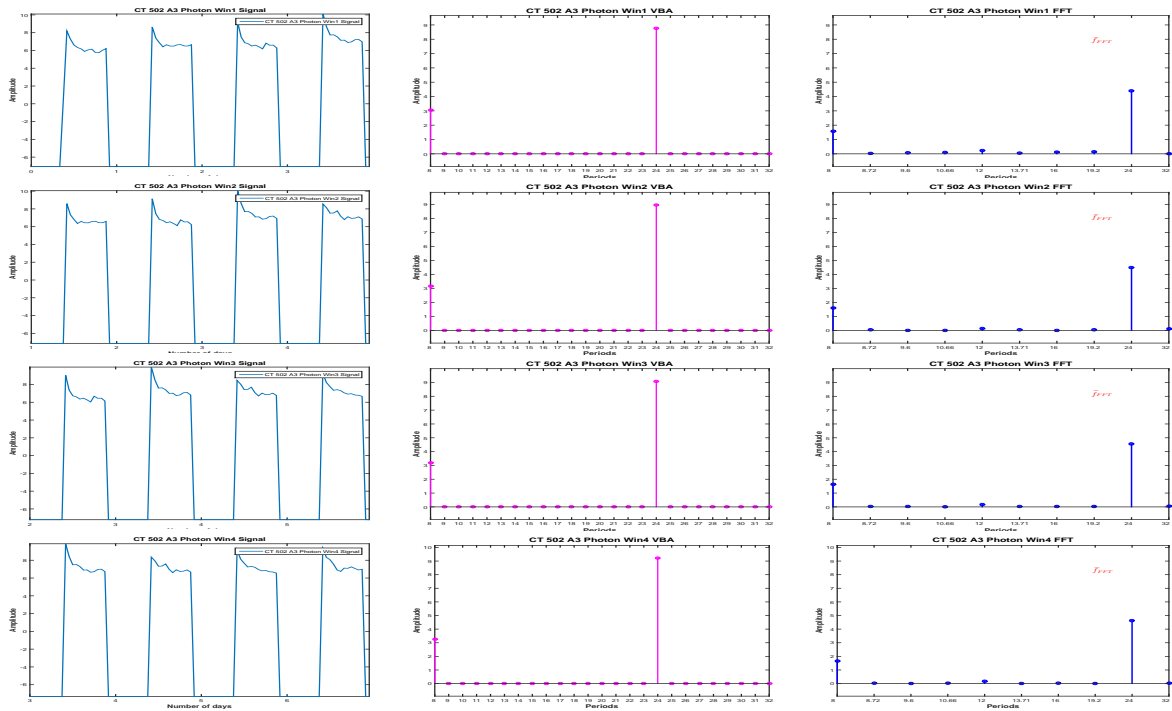


Figure 6.61: PC Stability: PC estimation via FFT and VBA for 4-days length signals

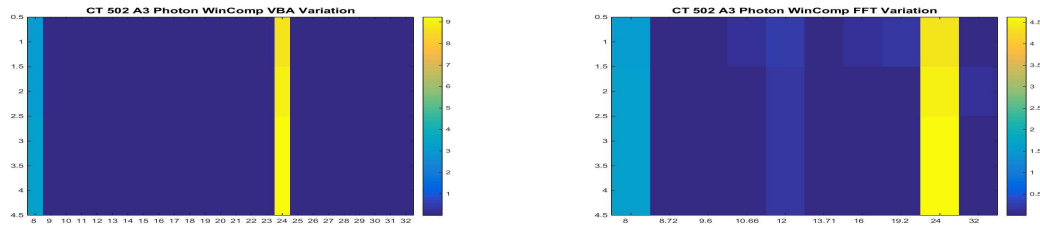


Figure 6.62: PC Stability: Proposed Method (a) vs. FFT (b)

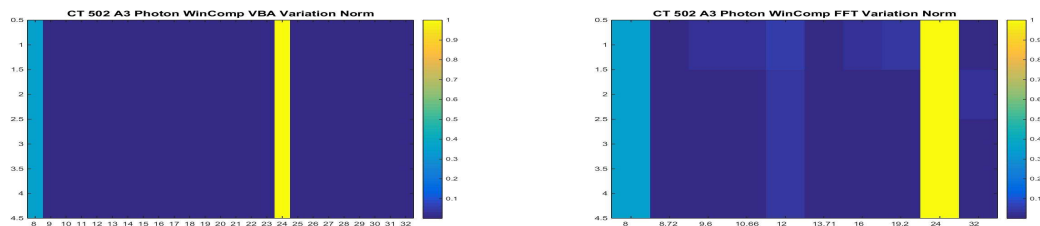


Figure 6.63: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

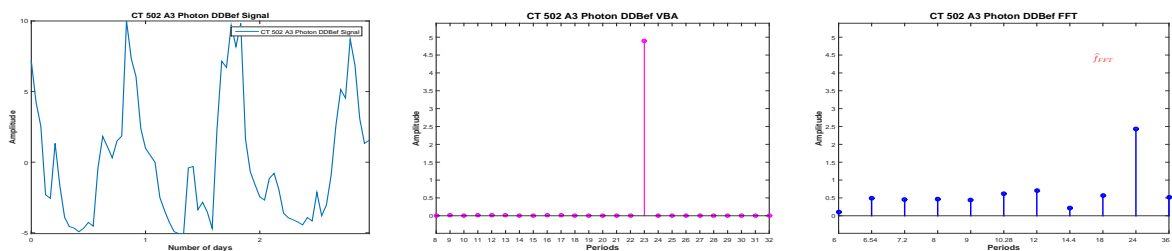


Figure 6.64: DD before signal (a) and the corresponding PC via VBA (b) and FFT (c)

Via FFT, we obtain a 24 hours periodicity, while via the Proposed Method, the dominant peak

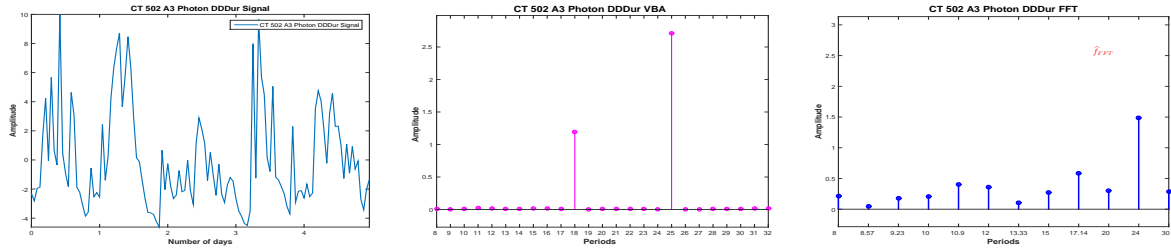


Figure 6.65: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

corresponds to 25. The only other non-zero amplitude in the PC amplitude vector, corresponds to 18 hour, which is consistent with the next highest peak in the FFT estimation, 17.14 hours. We note again the difference between the amplitudes corresponding to the two methods and the sparse PC vector returned by the Proposed Method.

We consider now 4-days length signals (windows) from the available **during** treatment part of the data, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.66).

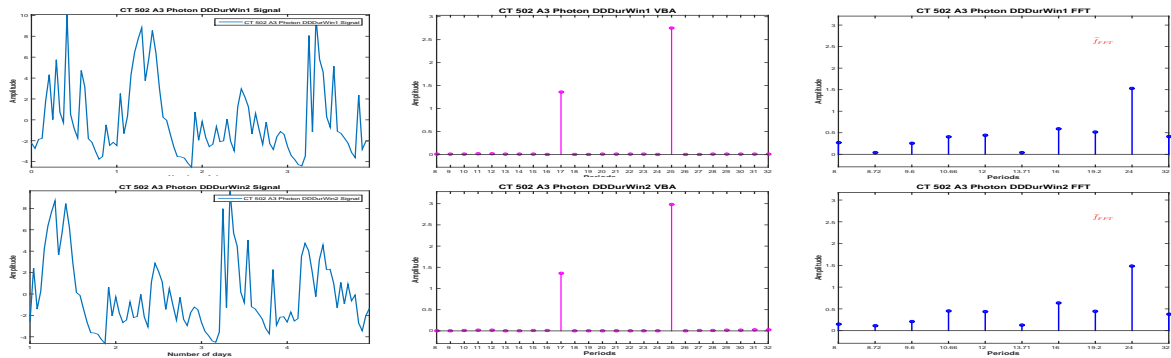


Figure 6.66: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during

Analysing the **during** segment using the moving window, a 24 hour periodicity is found for both windows, via FFT. However, via the Proposed Method, the periodicity, also stable, is estimated at 25 hours, Figure (6.67)

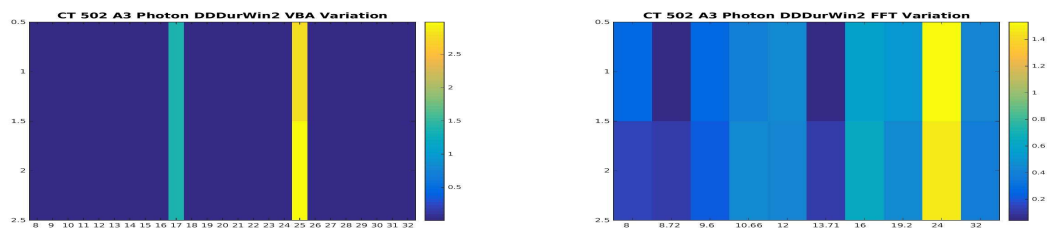


Figure 6.67: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.67):

### 6.3.3 - Conclusions: Dominant Period evolution

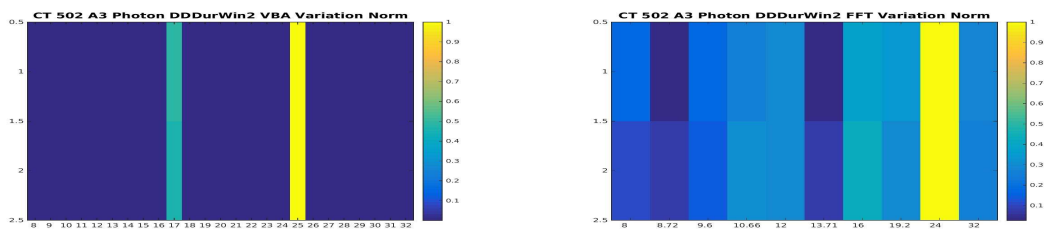


Figure 6.68: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

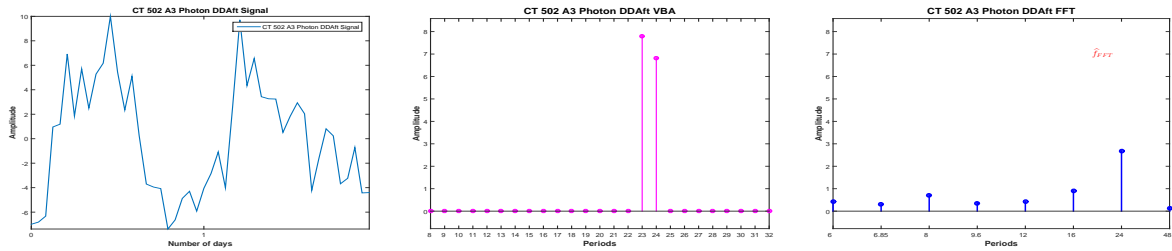


Figure 6.69: DD **after** signal (a) and the corresponding PC via VBA (b) and FFT (c)

We now consider the **after** treatment part of the data. We have a 2-days length signal. In this case, the dominant period is estimated at 24 hours via FFT. This value is very uncertain, for this particular case (a 2-days length signal) all the interval (16:48) having only one value that is appearing the FFT PC vector, i.e. 24 hours. Via the Proposed Method, the dominant period is set at 23 hours, with also an important periodicity for 24 hours.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Photon signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.6).

### 6.3.3 Conclusions: Dominant Period evolution

In this subsection we compared the FFT and Proposed Method PC amplitude vector evolution during the four segments of interest for mouse CT 503 A3, considering the gene expression (Photon Absorption signal) and the rest-activity patters (Activity signal), Figure (6.47). Like before, for both signals, four segments of interest were considered, i.e. the LD part of the signal, corresponding to the synchronization and the DD part of the signal, corresponding to the complete darkness, for which we have considered the segments corresponding to the **before**, **during** and **after** treatment phase, Figure (6.48) (rest-activity patterns) and Figure (6.59).

Like before, for the first segment of interest, i.e. the LD part of data and for the **during** treatment part, we have considered four days length signals to analyse the stability or variability of the dominant period for the considered segment itself.

The dominant period evolution, corresponding to the Proposed Method and FFT, for the gene expression data, is presented in Table (6.6). For this experiment, via FFT, the dominant period is 24 hours for each segment and seems stable. Via the Proposed Method, small variations are detected during the experiment: a 23 hour periodicity corresponding to the **before** and **after** treatment part of data and a 25 hour periodicity corresponding to the **during** treatment part of data.

CT 502 - A3	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	24	24
LD win 1	24	24
LD win 2	24	24
LD win 3	24	24
LD win 4	24	24
DD before	23	24
DD during	25	24
DD during win 1	25	24
DD during win 2	25	24
DD after	23	24

Table 6.6: CT 502 A3: Photon Dominant Period Stability: Proposed Method vs. FFT

The dominant period evolution, corresponding to the Proposed Method and FFT for the rest-activity patterns (Activity) signal is presented in Table (6.5). In this case, the result are similar, in the sense that both methods detect a circadian rhythm, but via the Proposed Method, variations of the dominant period can be detected, while via FFT the evolution of the dominant period seems stable.

## 6.4 CT 502 B1: Photon Absorption and Activity

In the following we consider the experiment **CT 502, Mouse B1**. This section is dedicated to the complete study, i.e the study of the physiology, corresponding to the segment LD (light-dark, 12h-12h), where the mouse was synchronized and DD (total darkness), the segment corresponding to the **before** treatment and the study of the behaviour **during** and **after** treatment. For the study, two signals were recorded: the **Activity Signal** measuring the activity (movement) of the mouse and the **Photon Absorption Signal**, measuring the number of photons absorbed, Figure (6.70).

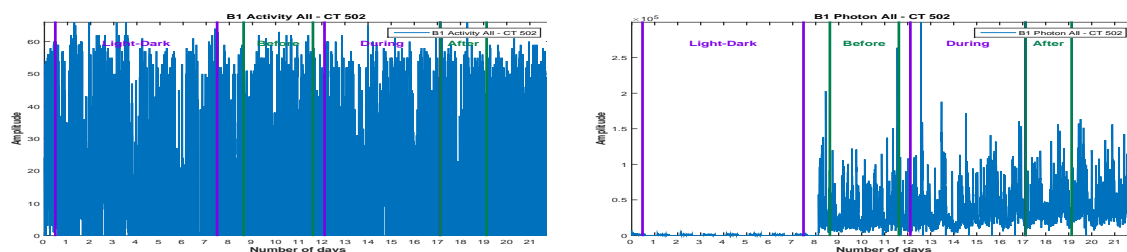


Figure 6.70: CT 502, B1 Mouse: Activity (a) and Photon Absorption (b) raw data

Figure (6.70) presents the raw data corresponding to the Activity signal and Photon Absorption signal. The signals were recorded for 22 days, sampled every minute. For each signal, we consider 4 segments: **LD Period**, corresponding to the synchronization part, then three parts corresponding to the **DD period**: before, during and after treatment. We are interested in the periodic component (PC) vector corresponding to each signal and in the corre-



### 6.4.1 - Dominant Period evolution: Activity

sponding acrophase. The stability of the period is verified using the moving window strategy, where the available data allows.

### 6.4.1 Dominant Period evolution: Activity

We start the analysis with the signal corresponding to the Activity. We consider zero-mean signals, normalized between [-10:10] and sampled every hour.

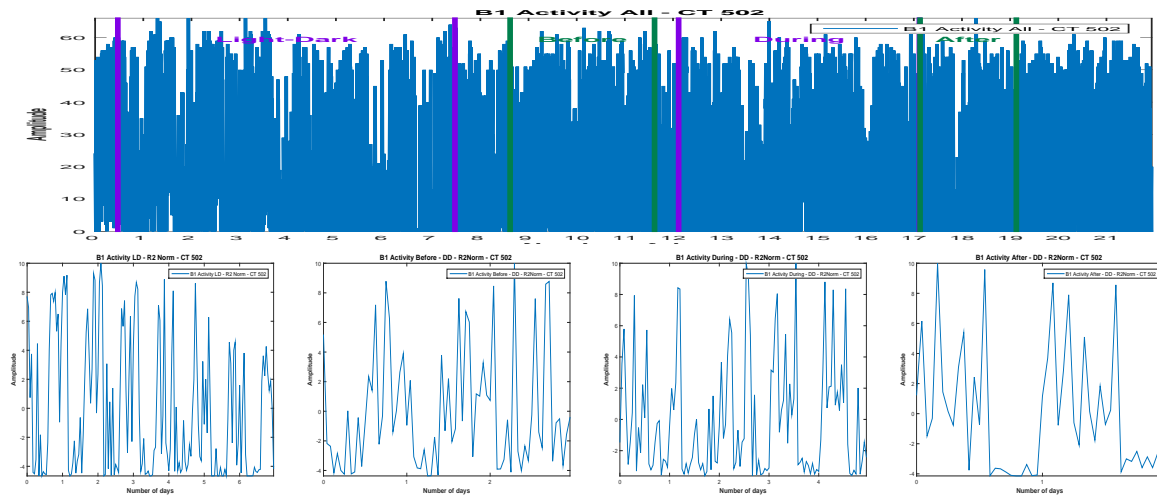


Figure 6.71: Activity raw data (a) and the corresponding parts, (b), (c), (d), normalized and one-hour sampled

For the LD part, 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 principal period

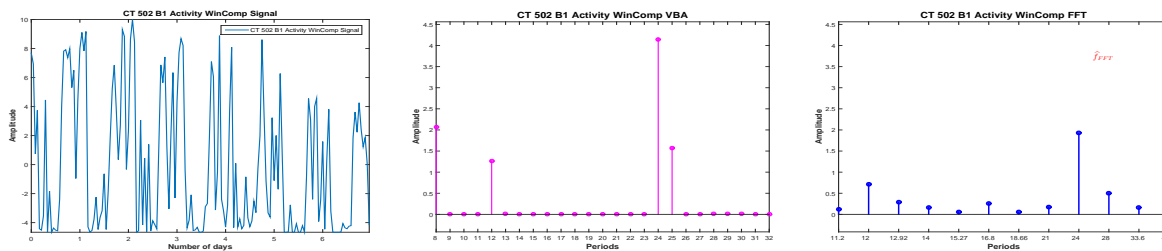


Figure 6.72: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

is estimated at 24 hours. Evidently, beside the incertitude associated with the FFT estimated PC, the existence of other rhythms can't be established, being difficult to interpret all the picks present in the estimated PC. Via the Proposed Method, the principal period is estimated at 24 hours. The amplitude corresponding to the 24h period via the Proposed Method is much higher.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method. Via the Proposed Method, the first windows presents a 25 hours dominant period, the second windows presents a 24 hours dominant period and for the last two windows, the dominant period is estimated at 23 hours. Via FFT method the principal period seems stable, every window presenting a dominant period of 24 hours.

## CT 502 B1: Photon Absorption and Activity

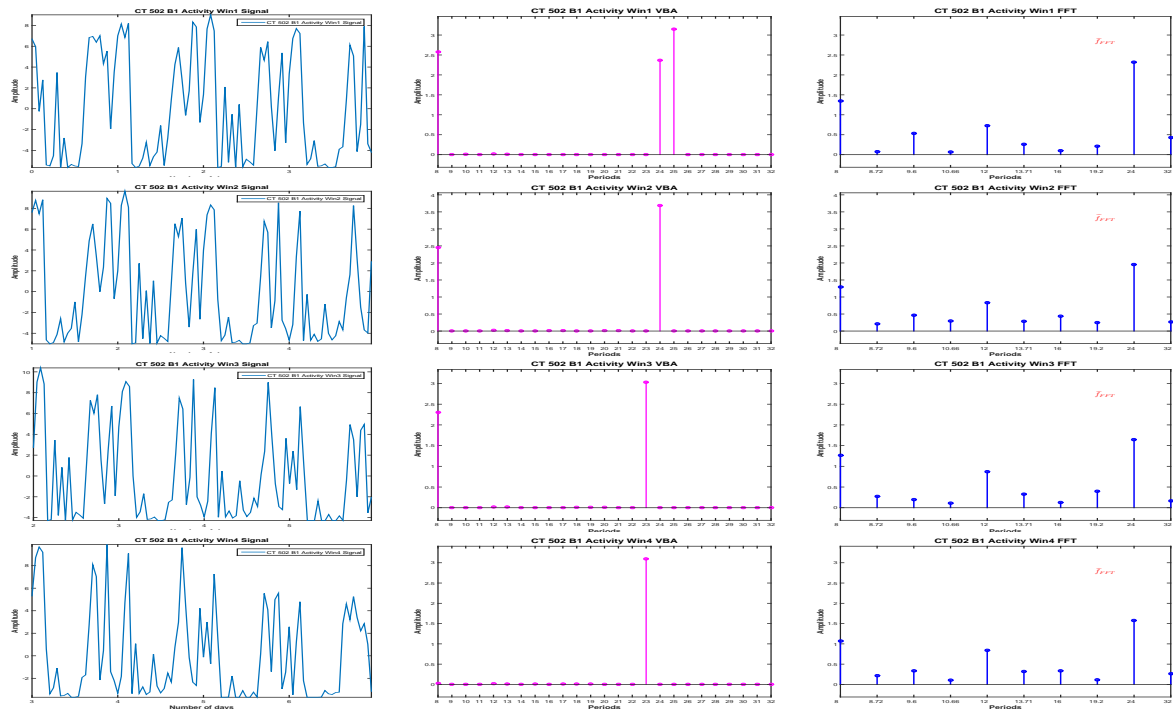


Figure 6.73: PC Stability: PC estimation via FFT and VBA for 4-days length signals

The stability or variability of the PC amplitudes vector for the LD part of the Activity signal is presented in Figure (6.74):

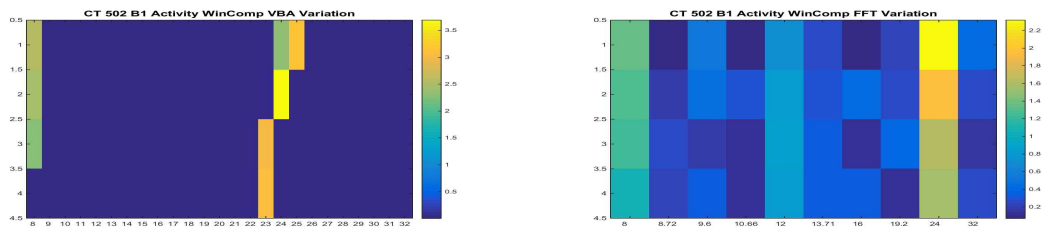


Figure 6.74: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in the Figure (6.75):

We consider now the **DD** period. First, we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, (6.76): Via the FFT, the highest pick is set at 24 hours and the next three highest picks are set at 8, 6 and 36 hours. Of course, 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 result presents a very important uncertainty. This means that for this particular case, even the classification of the rhythm (circadian or infradian) is uncertain via FFT. Via the Proposed method the dominant period is set at 23 hours.

We consider the **during** treatment part of the data, for which five days are available, Figure (6.77). 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 hours. For the FFT estimated PC vector, the dominant period is estimated at 24 hours, but the estimation is not sparse. In this

### 6.4.1 - Dominant Period evolution: Activity

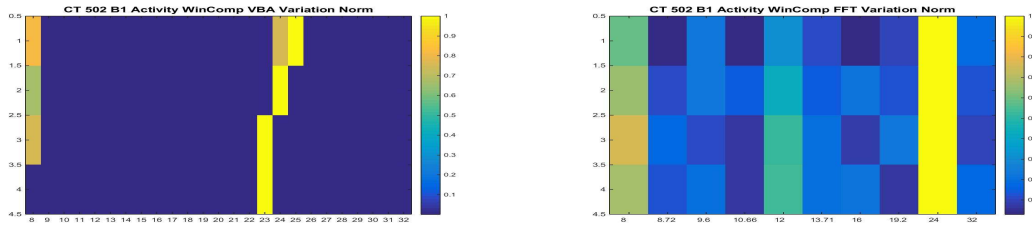


Figure 6.75: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

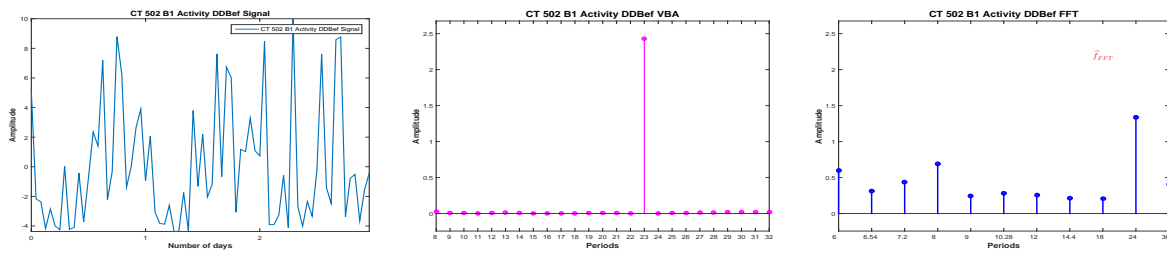


Figure 6.76: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

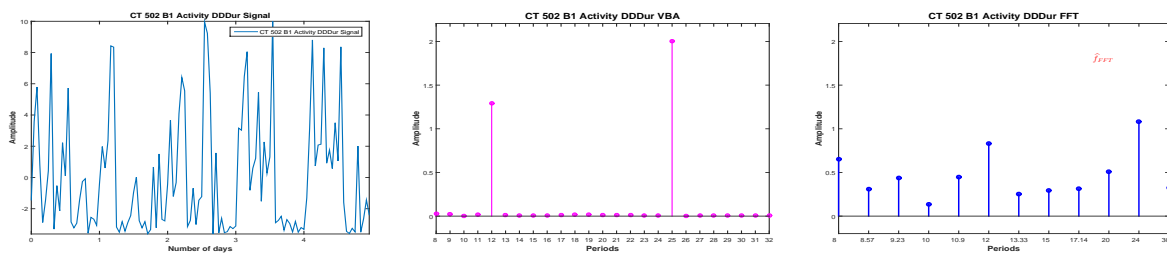


Figure 6.77: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

case, the next important pick corresponds to 12 hours, harmonics of 24 hours, and this values is also found in the estimation obtained by the Proposed Method.

We consider 4-days length signals (windows) from the **before** treatment segment, with a shift of one day and compute the PC via FFT and the Proposed method. In this case,

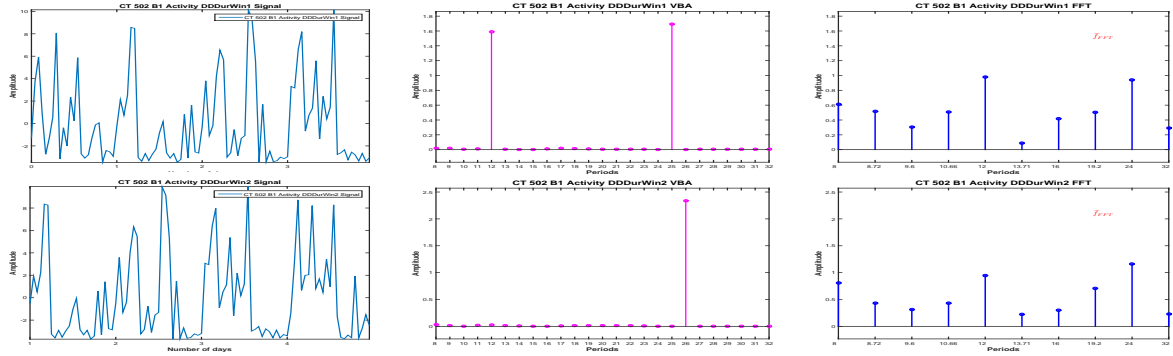


Figure 6.78: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during

analysing the **during** segment using the moving window we note that via the Proposed Method a variability of the dominant period is detected, 25 hours for the first window and 26 hours for the second. Via FFT the dominant period seems stable, 24 hours, but for the first window, the dominant period is 12 hours. This can be viewed in Figure (6.79). An image of dominant

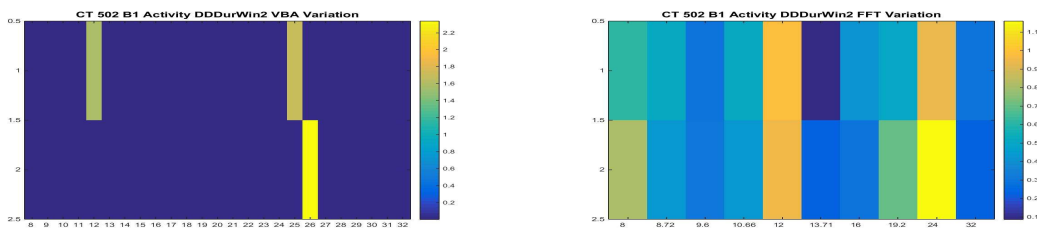


Figure 6.79: PC Stability: Proposed Method (a) vs. FFT (b)

period stability or variability, regardless the numerical value of the amplitude associated with the dominant period for **during** segment is presented in Figure (6.80):

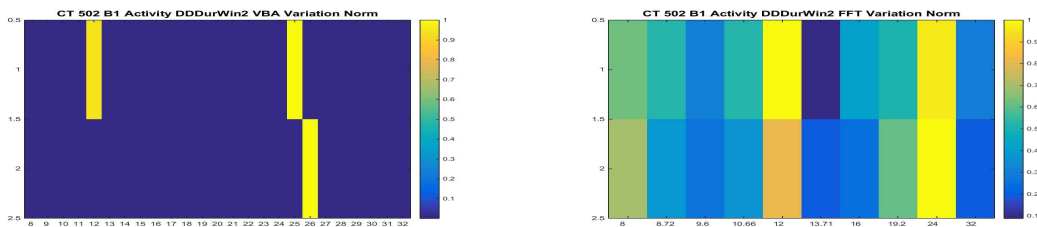


Figure 6.80: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

For the **after** treatment part of the data 2 days are available. Via the FFT estimation, the dominant period is 24 hours, while via the Proposed Method the dominant period is 26 hours.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Activity signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.7).

## 6.4.2 - Dominant Period evolution: Photon

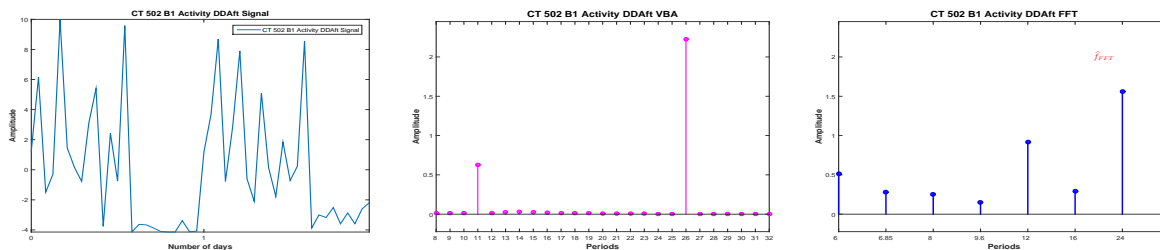


Figure 6.81: DD after signal (a) and the corresponding PC via VBA (b) and FFT (c)

CT 502 - B1	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	24	24
LD win 1	25	24
LD win 2	24	24
LD win 3	23	24
LD win 4	23	24
DD before	23	24
DD during	25	24
DD during win 1	25	12
DD during win 2	26	24
DD after	26	24

Table 6.7: CT 502 B1: Activity Dominant Period Stability: Proposed Method vs. FFT

## 6.4.2 Dominant Period evolution: Photon

In this section we analyse the signal corresponding to the photon absorption, i.e. the gene expression. Like before, we consider mean-zero signals, normalized between  $[-10:10]$  and sampled every hour. In this case, the analysis corresponding to the LD part has no biological interest, since the RT-BIO photomultiplier will not record the mouse liver bioluminescence but the light inside cage corresponding to the 12 hours of light.

The complete Photon absorption signal and the four segments of interest are presented in Figure (6.82).

For the LD part, 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 principal period is estimated at 24 hours. In particular, in this case, the FFT method estimates a sparse PC amplitudes vector, and along with the 24 hours, a non-zero peak is detected corresponding to 8 hours. Very similar results are obtained in this case via the Proposed Method: the principal period is also estimated at 24 hours, together with the harmonic of 8 hours. However, the amplitude corresponding to the Proposed Method is much higher.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.84). For all the windows, a 24 hours periodicity is confirmed, via the two methods, with a harmonic corresponding to 8 hours, Figure (6.85). Once again, for this segment, there is no biological interpretation.

# CT 502 B1: Photon Absorption and Activity

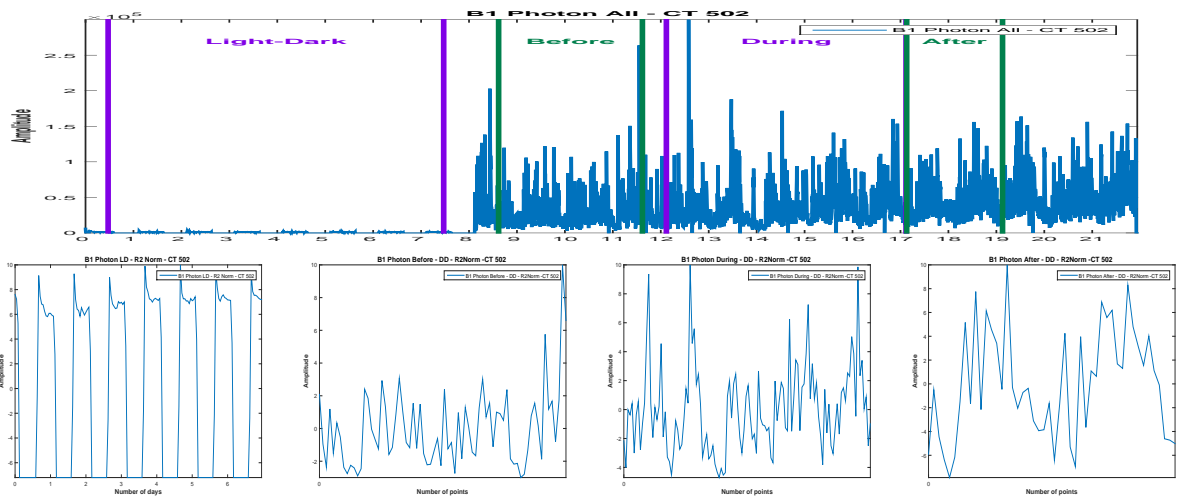


Figure 6.82: Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled

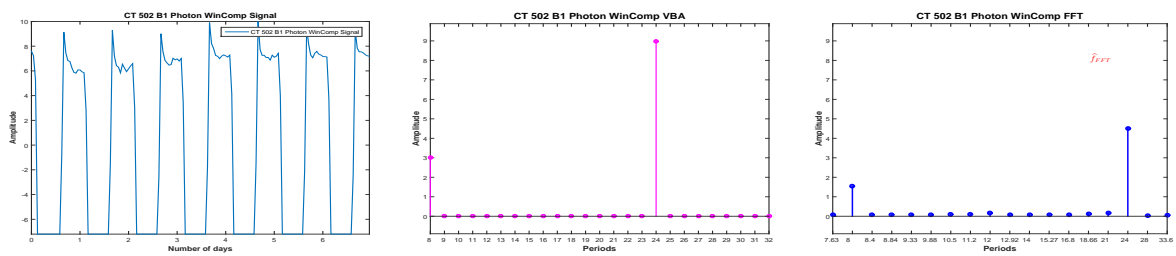


Figure 6.83: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

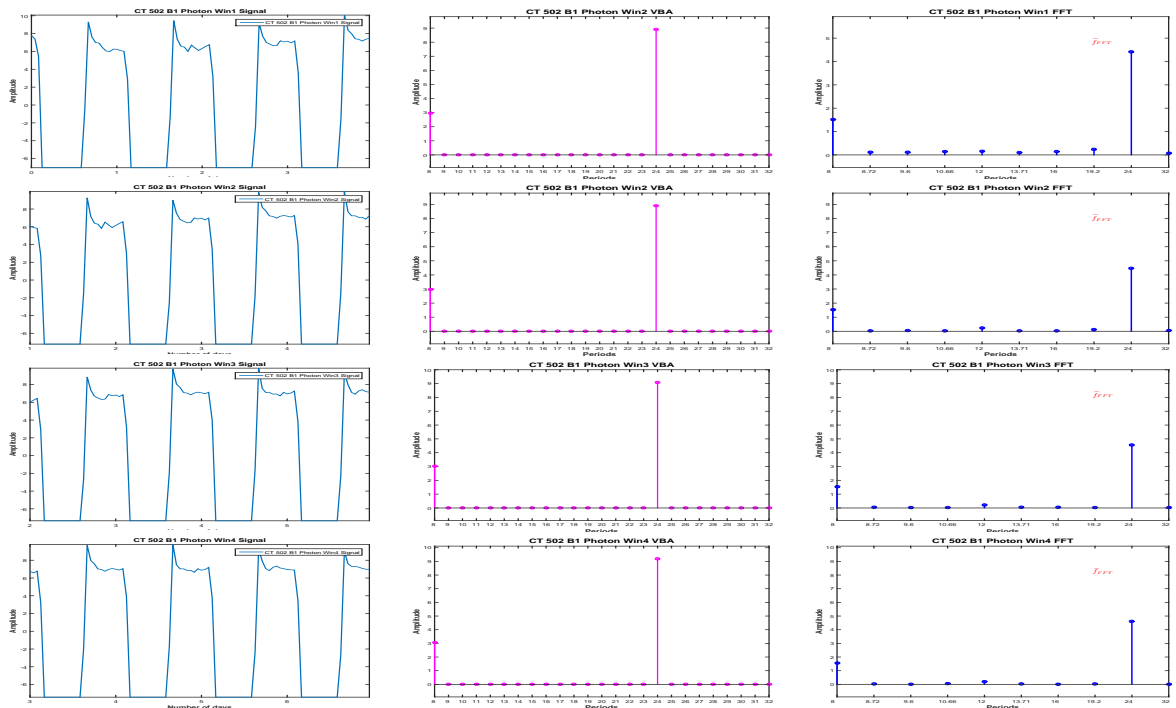


Figure 6.84: PC Stability: PC estimation via FFT and VBA for 4-days length signals

## 6.4.2 - Dominant Period evolution: Photon

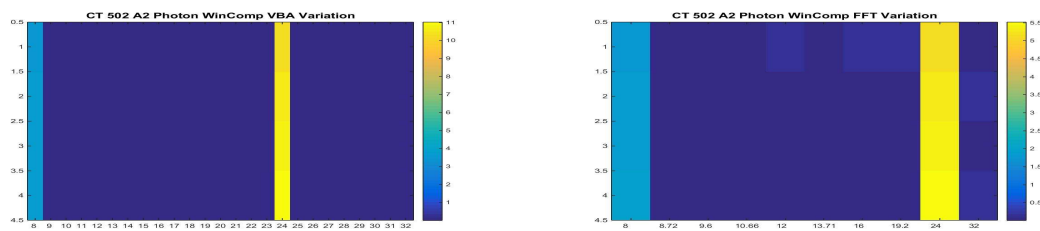


Figure 6.85: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.86):

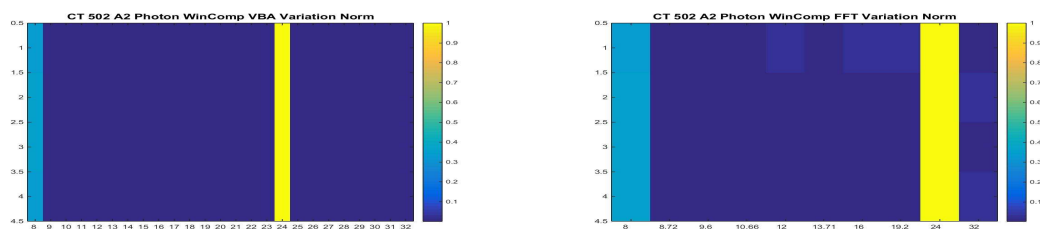


Figure 6.86: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

We consider now the DD period. First, we consider the **before** treatment segment, Figure (6.87): The dominant period corresponding to the FFT, is 24 hours, and via the Proposed

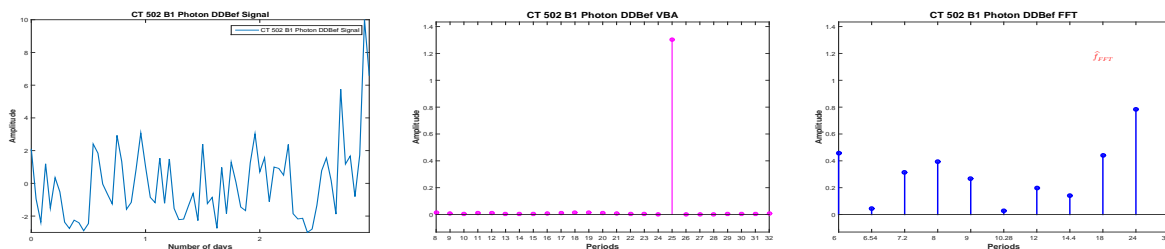


Figure 6.87: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

Method the estimated periodicity is 25 hours.

We consider the **during** treatment part of the data. Five days are available, Figure (6.88). Via FFT method, the dominant period is estimated at 24 hours, with important peaks corresponding to 20, 8 and 30 hours. Via the Proposed Method, the PC amplitudes vector detects three peaks: the dominant period is set at 25 hours and non-zero peaks are also detected for 19 and 20 hours.

The analysis of the **during** segment using the moving window is presented in Figure (6.89). For the FFT method, the dominant period corresponding to the first window is estimated at 24 hours, and for the second one is estimated at 19.2 hours. However in both cases, both peaks are almost equal. Via the Proposed Method, the dominant period corresponding to the first window is 25 hours and corresponding to the second window is 26 hours. For both windows, an important peak is estimated at 19 hours, consistent with the FFT estimation.

## CT 502 B1: Photon Absorption and Activity

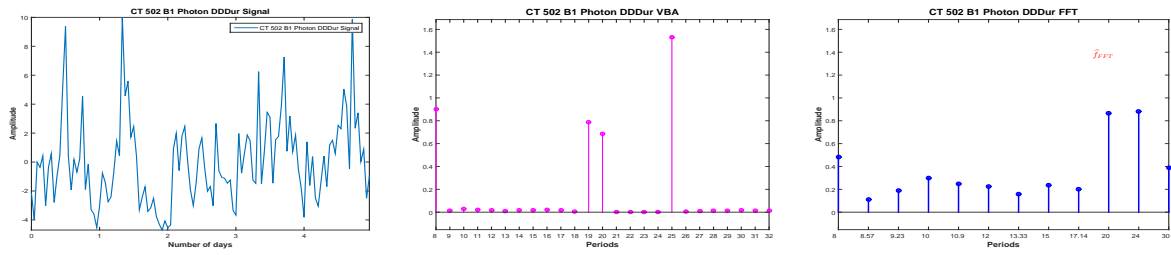


Figure 6.88: DD during signal (a) and the corresponding PC via VBA (b) and FFT (c)

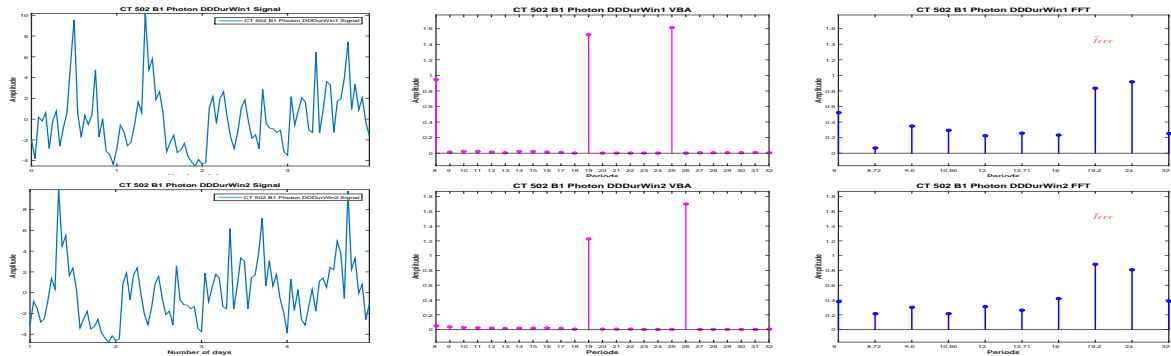


Figure 6.89: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during

The visual representation of the variability of the PC amplitude vector via the two methods is presented in Figure (6.90).

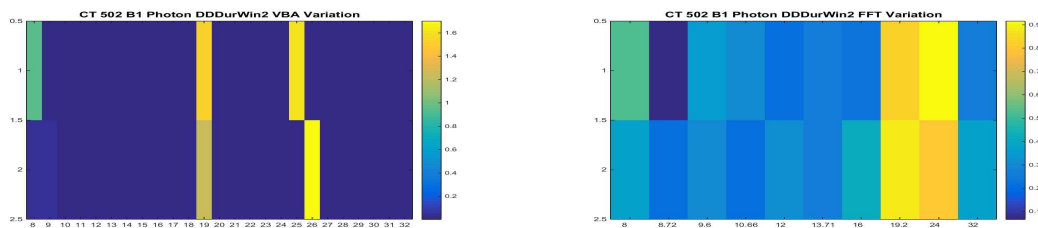


Figure 6.90: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.91).

We consider the **after** treatment part of the data, two days length signal and compute the PC corresponding to the **after** treatment part of the data using the Proposed Method and also using the FFT, Figure (6.92). In this case, the dominant period is estimated at 24 hours both via FFT and 23 hours corresponding to the Proposed Method.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Photon absorption signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.8).



## 6.4.2 - Dominant Period evolution: Photon

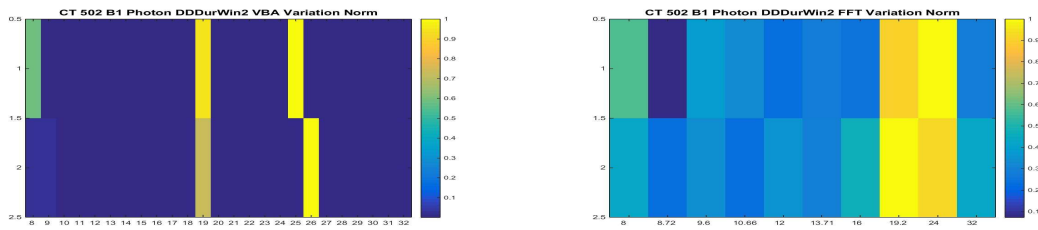


Figure 6.91: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

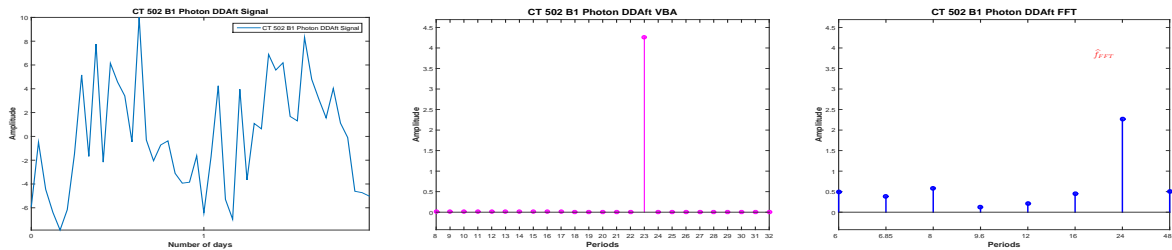


Figure 6.92: DD **after** signal (a) and the corresponding PC via VBA (b) and FFT (c)

CT 502 - B1	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	24	24
LD win 1	24	24
LD win 2	24	24
LD win 3	24	24
LD win 4	24	24
DD before	25	24
DD during	25	24
DD during win 1	25	24
DD during win 2	26	19
DD after	23	24

Table 6.8: CT 502 B1: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT

### 6.4.3 Conclusions: Dominant Period evolution

We compared the FFT and Proposed Method PC amplitude vector evolution during the four segments of interest for mouse CT 503 B1, considering the gene expression (Photon Absorption signal) and the rest-activity patterns (Activity signal), Figure (6.70). For both data, we have considered four segments of interest, i.e. the LD part of the signal, corresponding to the synchronization and the DD part of the signal, corresponding to the complete darkness, for which we have considered the segments corresponding to the **before**, **during** and **after** treatment phase, Figure (6.71) (rest-activity patterns) and Figure (6.82).

For the first segment of interest, i.e. the LD part of data and for the **during** treatment part, we have considered four days length signals to analyse the stability or variability of the dominant period for the considered segment itself.

The dominant period evolution, corresponding to the Proposed Method and FFT for the gene expression data (photon absorption signal), is presented in Table (6.8). For this experiment, via FFT, the dominant period is 24 hours for all segments, except the second window of the **during** treatment segment and seems stable. Via the Proposed Method, slight variations are detected during the experiment: a 25 hour periodicity corresponding to the **before** and **during** treatment part of data and a 23 hour periodicity corresponding to the **after** treatment part of data.

In the study of the stability of the dominant period for the **during** treatment segment, Figure (6.78), for both windows, via both methods, important peaks are estimated for 19 and 24 hours (25 for the second one, via Proposed Method). Via FFT, for the second window, the dominant period corresponds to 19 hours. Via Proposed Method, for the second window, the dominant period corresponds to 25 hours.

The dominant period evolution, corresponding to the Proposed Method and FFT for the rest-activity patterns (Activity) signal is presented in Table (6.7). Via FFT, the dominant period is 24 hours for all segments, except the first window of the **during** treatment segment and seems stable. Via the Proposed Method, slight variations around 24 hours are detected during the experiment: a 23 hour periodicity corresponding to the **before** treatment part of data, a 25 hour periodicity corresponding to the **during** treatment part of data and a 26 hour periodicity corresponding to the **after** treatment part of data.

In the study of the stability of the dominant period for the **during** treatment segment, Figure (6.89), for both windows, via FFT, important peaks are estimated for 12 and 24 hours. For the first window, the dominant period corresponds to 12 hours. However, in this case, a circadian rhythm can be associated.

## 6.5 CT 502 B2: Photon Absorption and Activity

---

This section presents the comparison between the Proposed Method and the FFT method, for another experiment, performed in the same conditions as the one presented in Section (6.1).

For the study, two signals are recorded: the **Activity Signal** measuring the activity (movement) of the mouse and the **Photon Absorption Signal**, measuring the number of photon absorbed. Figure (6.93) presents the raw data corresponding to the Activity signal and Photon Absorption signal. The signals were recorded for 22 days, sampled every minute. Like in

### 6.5.1 - Dominant Period evolution: Activity

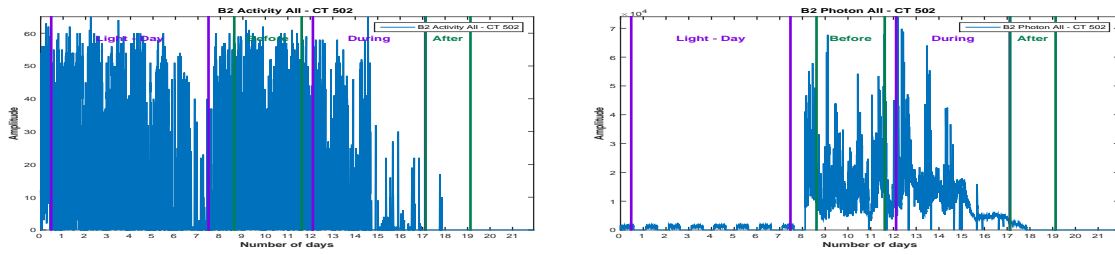


Figure 6.93: CT 502, B2 Mouse: Activity (a) and Photon Absorption (b) raw data

Section (6.1), for each signal, we consider the same 4 segments: **LD Period**, corresponding to the synchronization part, then three parts corresponding to the **DD period**: before, during and after treatment. We are interested in the periodic component (PC) vector corresponding to each signal. The stability of the period is verified using the moving window strategy, where the available data allows.

### 6.5.1 Dominant Period evolution: Activity

We present the analysis with the signal corresponding to the Activity. We will consider mean-zero signals, normalized between  $[-10:10]$  and sampled every hour, Figure (6.94).

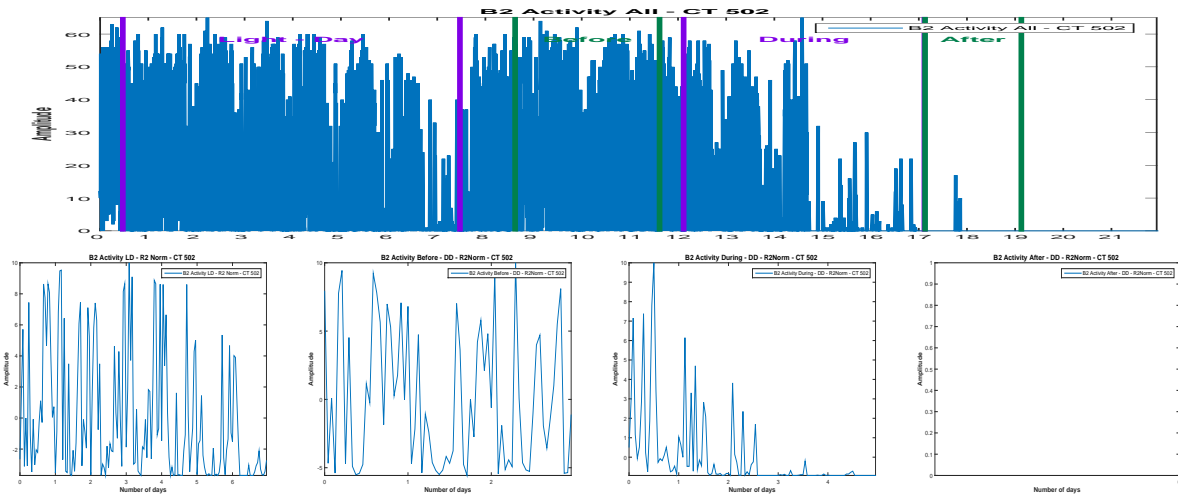


Figure 6.94: Activity raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled

For the LD part, 7 days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.95): Via the FFT method, the principal period is estimated at 24 hours and the next highest peak in the PC amplitude vector corresponds to 12 hours. Via the Proposed Method, the principal period is estimated at 25 hours, together with the harmonic of 12 hours. However, the amplitude corresponding to the Proposed Method is much higher.

We consider four days length signals (windows) from the available LD signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.96).

Via FFT method, the dominant period is 24 hours for all windows, therefore stable. Via the Proposed Method, the dominant period is not stable, corresponding to 25 hours for the first

## CT 502 B2: Photon Absorption and Activity

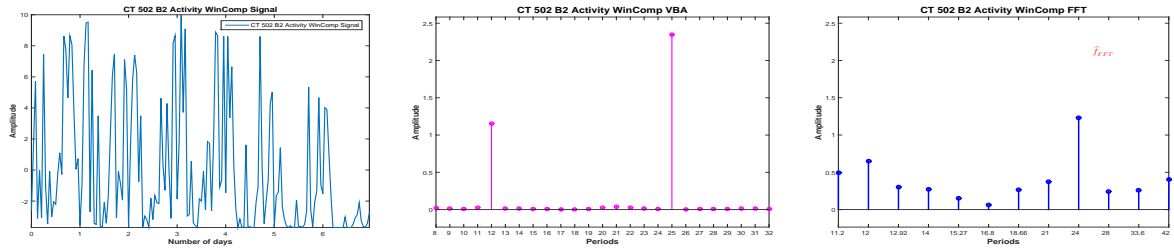


Figure 6.95: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

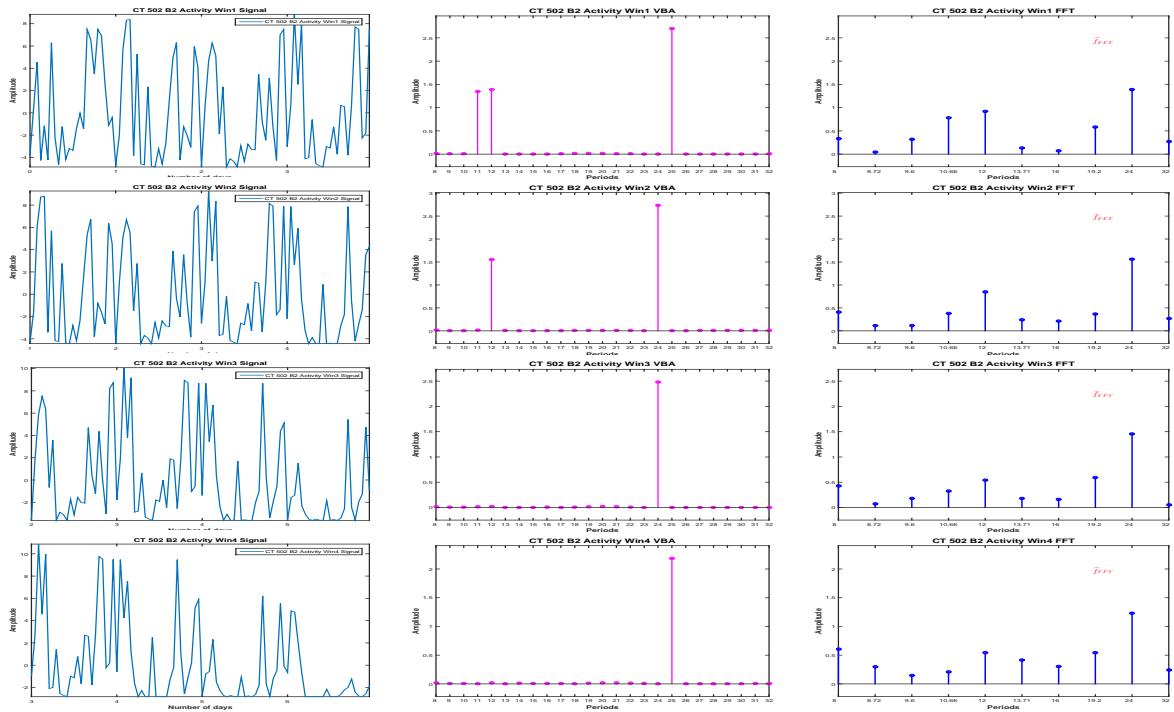


Figure 6.96: PC Stability: PC estimation via FFT and VBA for 4-days length signals

## 6.5.2 - Dominant Period evolution: Photon

window, 24 hours for the next two windows and 25 hours for the last window, Figure (6.97).

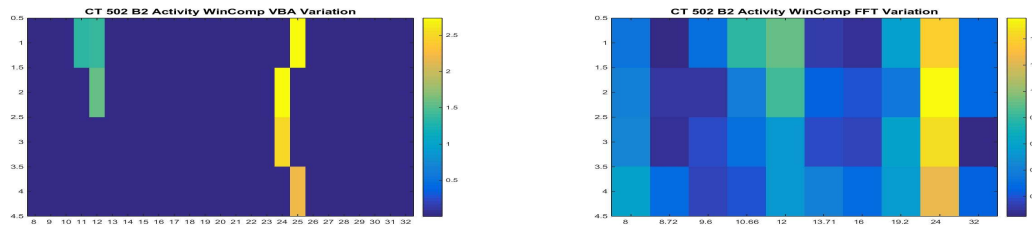


Figure 6.97: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.98):

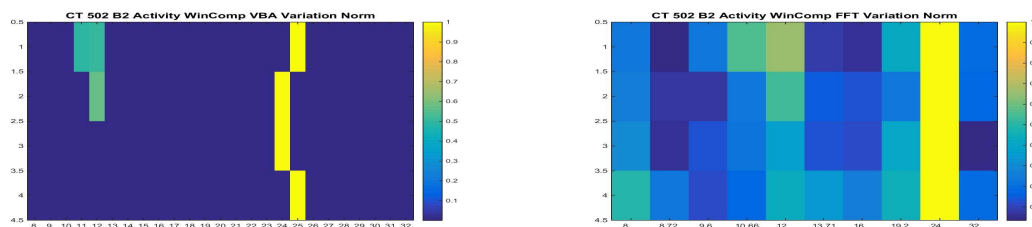


Figure 6.98: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

For the DD period, we consider first the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, Figure (6.99): The estimation corresponding to the FFT method, the dominant peak corresponds

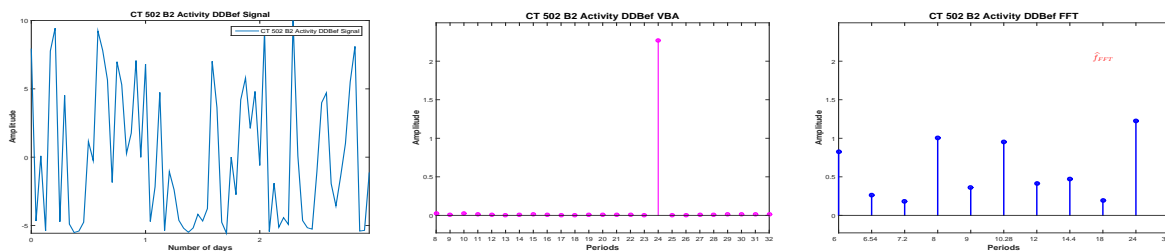


Figure 6.99: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

to 24 hours. For the estimation corresponding to the Proposed Method, the PC amplitude vector is sparse, with only one non-zero peak, corresponding to 24.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Activity signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.9).

## 6.5.2 Dominant Period evolution: Photon

In this section we analyse the signal corresponding to the photon absorption, i.e. the gene expression. Like before, we consider mean-zero signals, normalized between [-10:10] and

CT 502 - B2	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	23	24
LD win 1	21	24
LD win 2	22	24
LD win 3	23	24
LD win 4	25	24
DD before	24	24
DD during	-	-
DD during win 1	-	-
DD during win 2	-	-
DD after	-	-

Table 6.9: CT 502 B2: Activity Dominant Period Stability: Proposed Method vs. FFT

sampled every hour. In this case, the analysis corresponding to the LD part has no biological interest, since the RT-BIO photomultiplier will not record the mouse liver bioluminescence but the light inside cage corresponding to the 12 hours of light. The complete Photon absorption signal and the four segments of interest are presented in Figure (6.100).

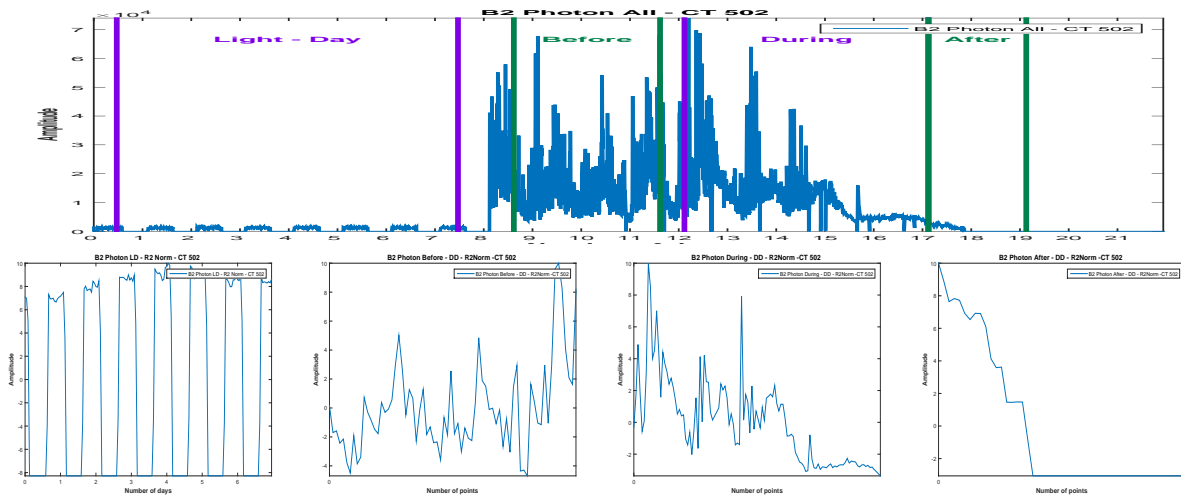


Figure 6.100: Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled

For the LD part, 8 days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.101): Via the FFT method, the principal period is estimated at 24 hours and the next highest peak in the PC amplitude vector corresponds to 12 hours. Via the Proposed Method, the principal period is estimated at 25 hours, together with the harmonic of 12 hours. However, the amplitude corresponding to the Proposed Method is much higher.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.102).

For all the windows, a 24 hours periodicity is confirmed, via the two methods, Fig-

## 6.5.2 - Dominant Period evolution: Photon

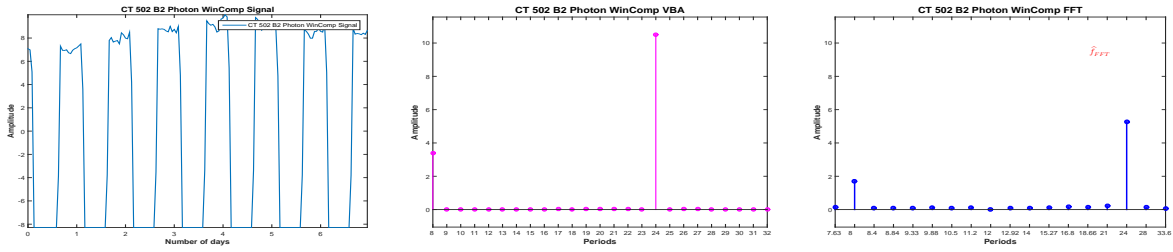


Figure 6.101: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

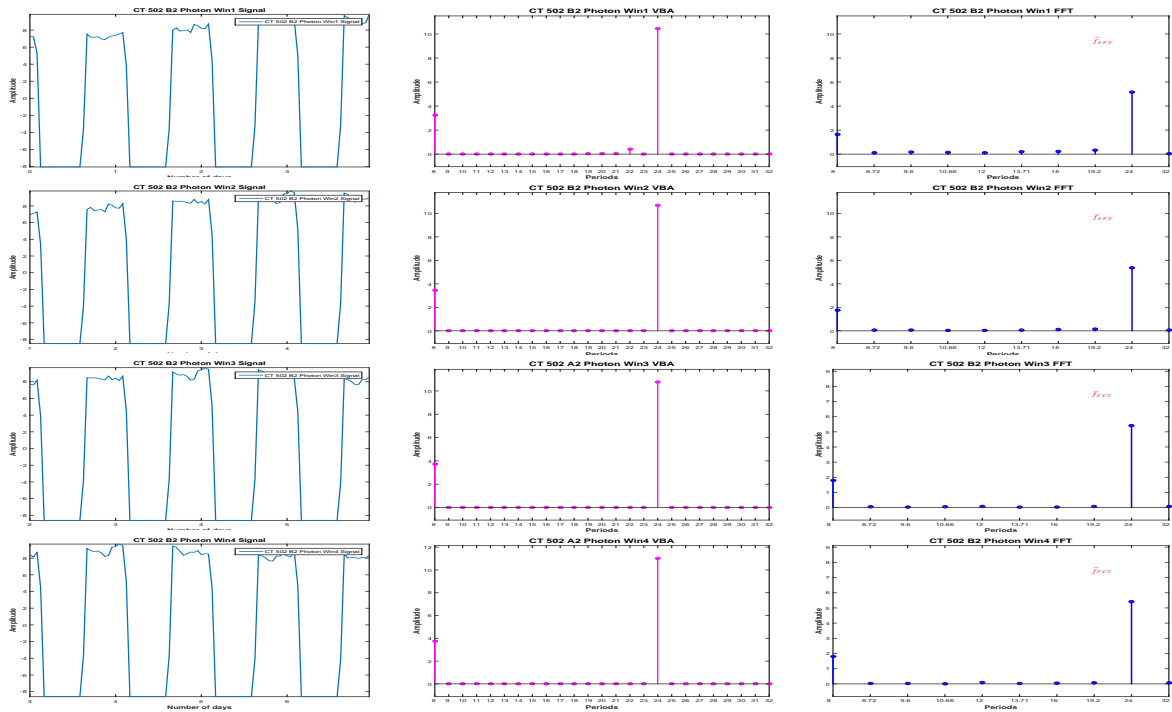


Figure 6.102: PC Stability: PC estimation via FFT and VBA for 4-days length signals

ure (6.103).

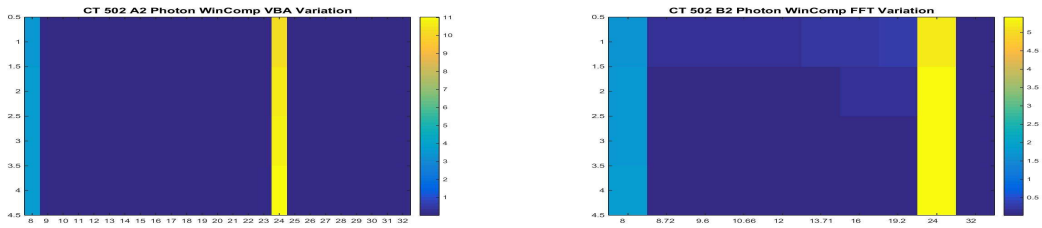


Figure 6.103: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.104):

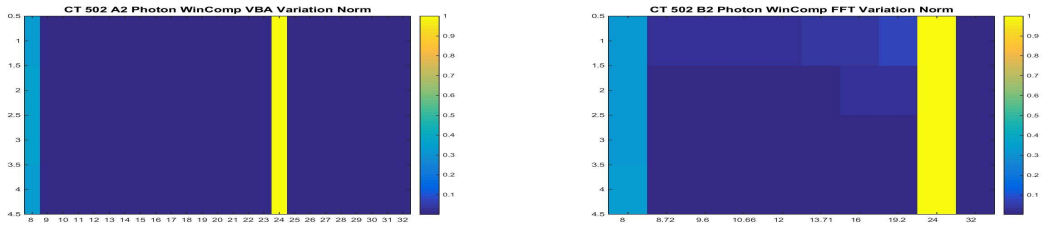


Figure 6.104: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

For the DD period, first we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, Figure (6.105):

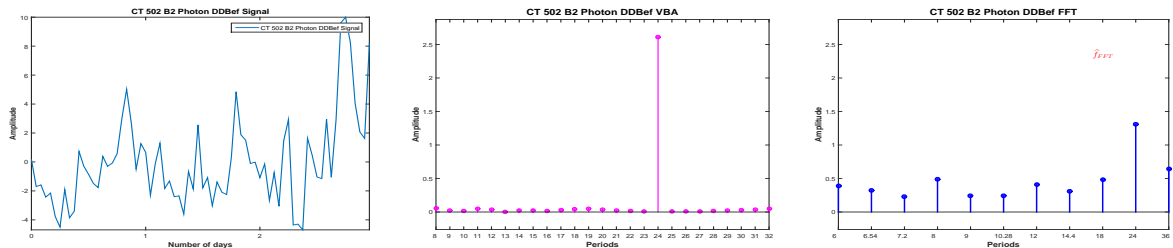


Figure 6.105: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

Via the FFT, the highest pick is set at 24. Of course, 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 25 hours.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Photon absorption signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.10).



## 6.5.2 - Dominant Period evolution: Photon

CT 502 - B2	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	24	24
LD win 1	24	24
LD win 2	24	24
LD win 3	24	24
LD win 4	24	24
DD before	25	24
DD during	24	24
DD during win 1	24	24
DD during win 2	24	24
DD after	23	24

Table 6.10: CT 502 B2: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT

## 6.6 CT 502 B3: Photon Absorption and Activity

In the following we consider the experiment **CT 502, Mouse B3**. This section is dedicated to the complete study, i.e the study of the physiology, corresponding to the segment LD (light-dark, 12h-12h), where the mouse was synchronized and DD (total darkness), the segment corresponding to the **before** treatment and the study of the behaviour **during** and **after** treatment. For the study, two signals were recorded: the **Activity Signal** measuring the activity (movement) of the mouse and the **Photon Absorption Signal**, measuring the number of photon absorbed, Figure (6.106).

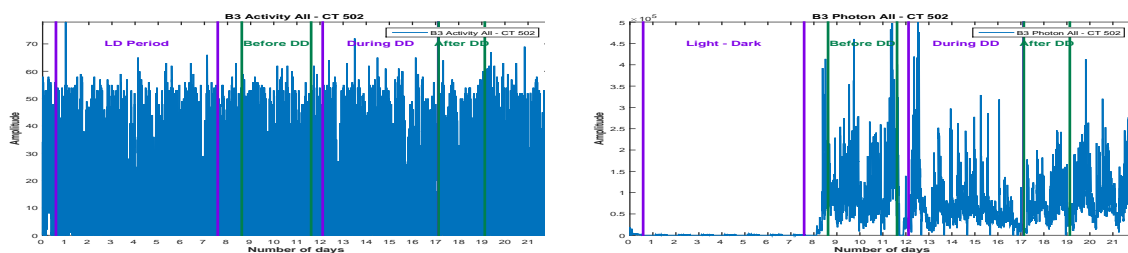


Figure 6.106: CT 502, B3 Mouse: Activity (a) and Photon Absorption (b) raw data

Figure (6.106) presents the raw data corresponding to the Activity signal and Photon Absorption signal. The signals were recorded for 22 days, sampled every minute. For each signal, we consider 4 segments: **LD Period**, corresponding to the synchronization part, then three parts corresponding to the **DD period**: before, during and after treatment. We are interested in the periodic component (PC) vector corresponding to each signal and in the corresponding acrophase. The stability of the period is verified using the moving window strategy, where the available data allows.

### 6.6.1 Dominant Period evolution: Activity

We start the analysis with the signal corresponding to the Activity. We will consider mean-zero signals, normalized between [-10:10] and sampled every hour. The Activity signal and the four segments of interest are presented in Figure (6.107).

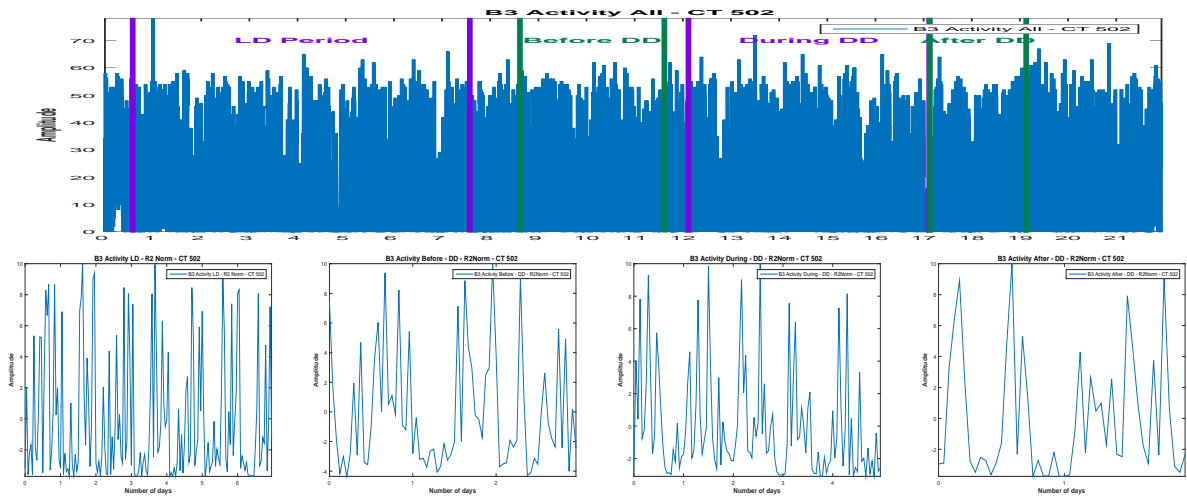


Figure 6.107: Activity raw data (a) and the corresponding parts, (b), (c), (d), normalized and one-hour sampled

For the LD part, we have 8 days available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.108). Via the FFT method,

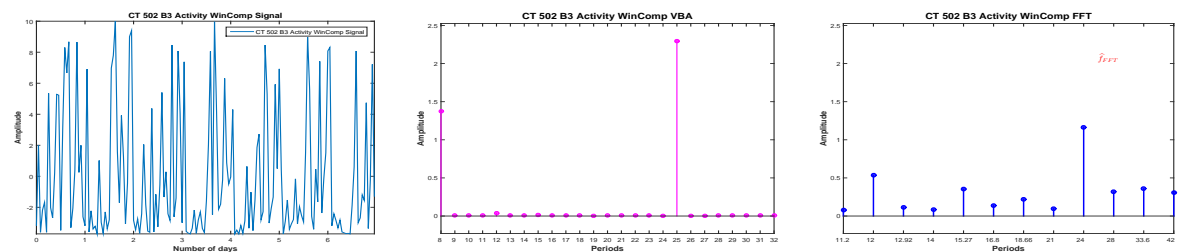


Figure 6.108: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

the principal period is estimated at 24 hours. The next highest pick in the PC vector is corresponding to 12 hours. Via the Proposed Method, the principal period is estimated at 25 hours.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.109).

Via FFT, the dominant period seems stable, for each window the dominant period being estimated at 24 hours. Via the Proposed Method, the dominant period shifts from 25 hours, for the first window to 24 hours for the second window and third window to 25 hours for the last window, Figure (6.110).

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented Figure (6.111):.

We consider now the DD period. First, we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed

## 6.6.1 - Dominant Period evolution: Activity

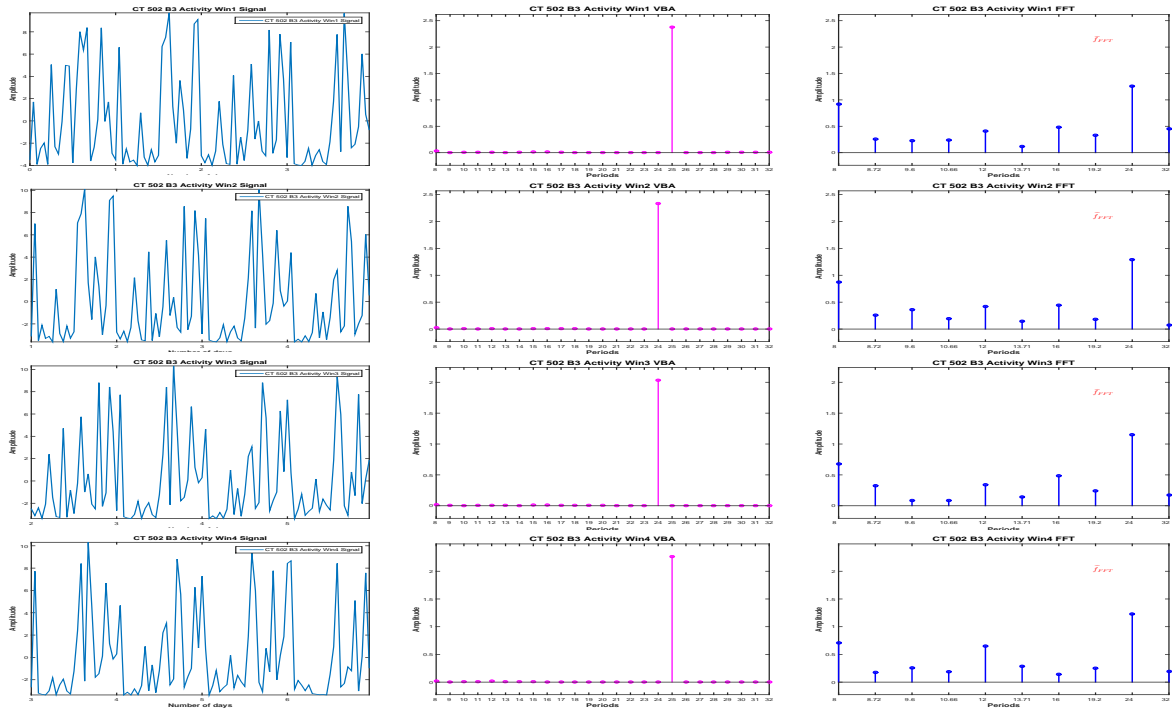


Figure 6.109: PC Stability: PC estimation via FFT and VBA for 4-days length signals

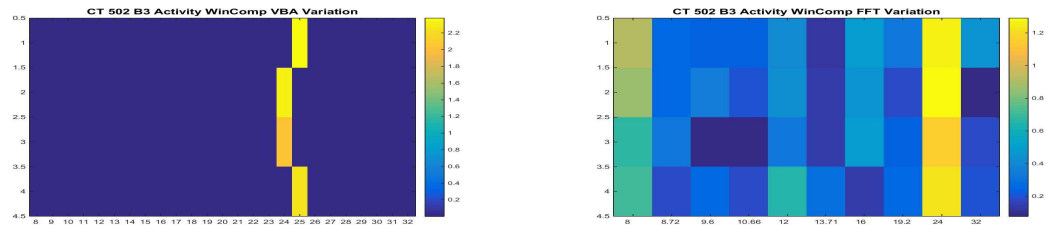


Figure 6.110: PC Stability: Proposed Method (a) vs. FFT (b)

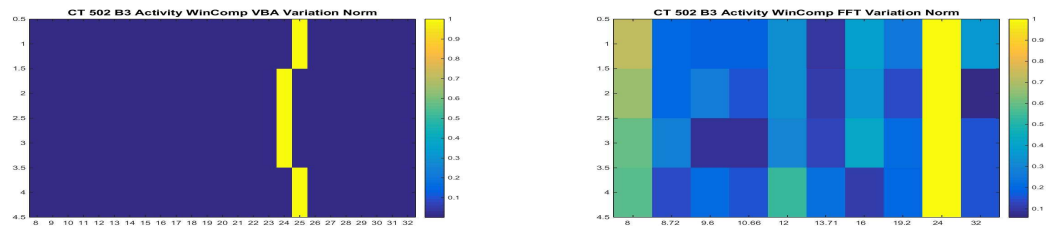


Figure 6.111: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

Method, Figure (6.112). Via the FFT, the highest pick is set at 24 hours and the next highest

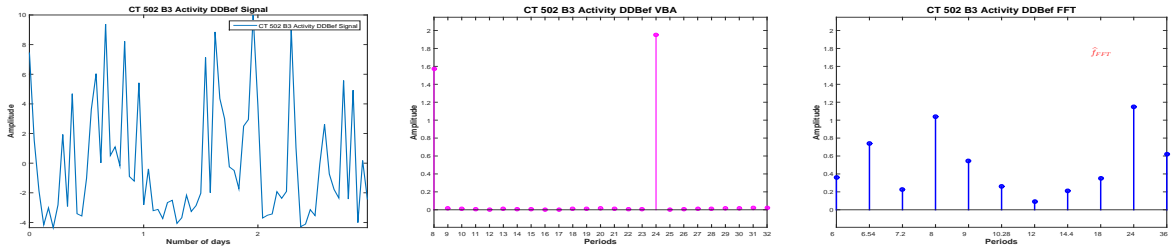


Figure 6.112: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

pick is set at 8 hours. Via the Proposed method, the PC amplitude vector is sparse, with only one non-zero amplitude, corresponding to 24 hours.

For the **during** treatment part of the data, 5 days are available. We compute the PC corresponding to the **during** treatment part of the data using the Proposed Method and also using the FFT, Figure (6.113). Via the Proposed Method, the estimated PC amplitude vector is a

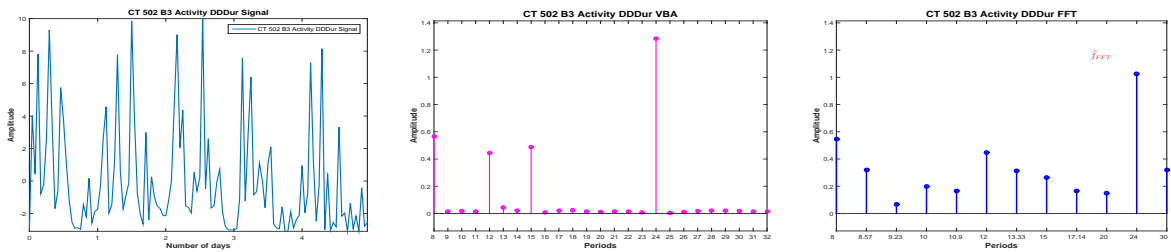


Figure 6.113: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

sparse vector and the dominant period is estimated at 24 hours. The other non-zero amplitudes in the PC amplitudes vector corresponds to 8, 12 and 15 hours. In the FFT estimation, the dominant period corresponds to 24 hours, and the next two highest peaks in the PC amplitude vector estimation corresponds to 8 and 12 hours.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.114).

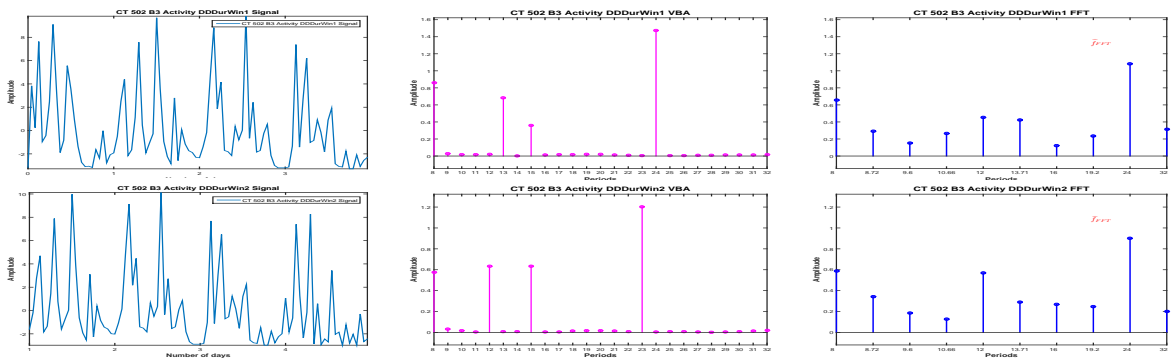


Figure 6.114: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during

We note that via the FFT, both windows are estimated with a dominant period of 24 hours, with harmonics corresponding to 8 and 12 hours. Via the Proposed Method, we find the

## 6.6.2 - Dominant Period evolution: Photon

dominant period corresponding to the first window at 24 hours (and important picks for 8 and 13 hours) and the dominant period corresponding to the second window at 23 hours (and important picks for 8 and 13 hours), Figure (6.115).

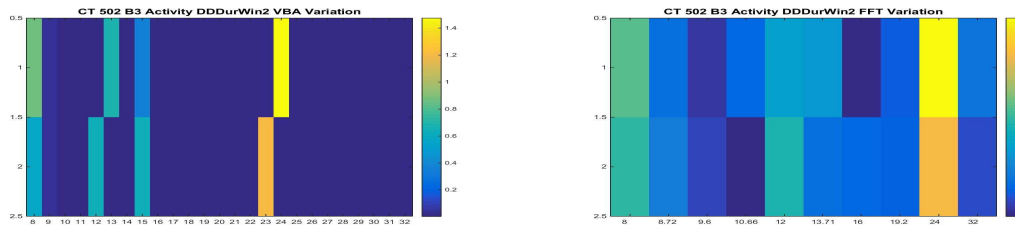


Figure 6.115: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in the following figure:

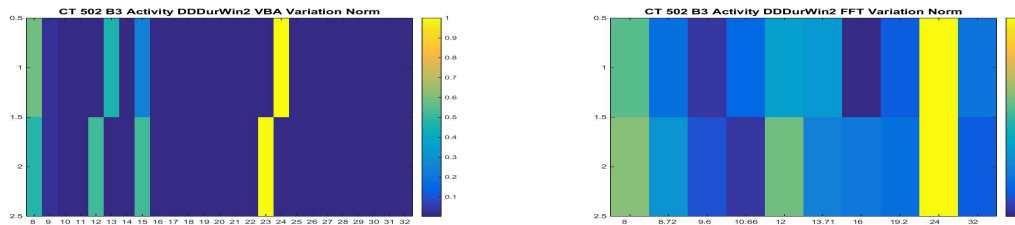


Figure 6.116: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

We now consider the **after** treatment part of the data. Two days are available. For the FFT,

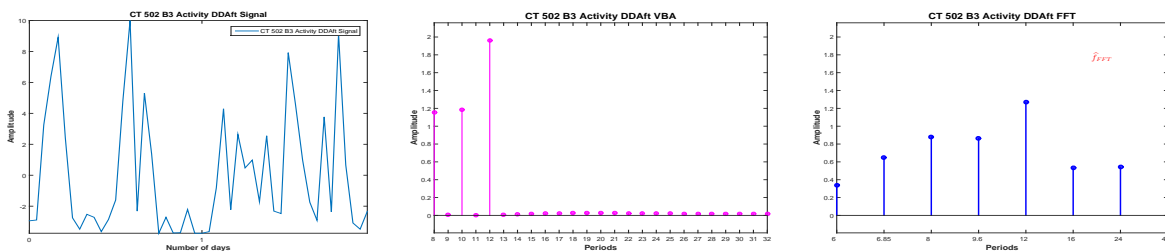


Figure 6.117: DD **after** signal (a) and the corresponding PC via VBA (b) and FFT (c)

the dominant period corresponds to 12 hours. Via the Proposed Method, the dominant period is also estimated at 12 hours.

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Activity signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.11).

## 6.6.2 Dominant Period evolution: Photon

In this section we analyse the signal corresponding to the photon absorption, i.e. the gene expression. Like before, we consider mean-zero signals, normalized between [-10:10] and

CT 502 - B3	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	25	24
LD win 1	25	24
LD win 2	24	24
LD win 3	24	24
LD win 4	25	24
DD before	24	24
DD during	24	24
DD during win 1	24	24
DD during win 2	23	24
DD after	12	12

Table 6.11: CT 502 B3: Activity Dominant Period Stability: Proposed Method vs. FFT

sampled every hour. In this case, the analysis corresponding to the LD part has no biological interest, since the RT-BIO photomultiplier will not record the mouse liver bioluminescence but the light inside cage corresponding to the 12 hours of light. The complete Photon absorption signal and the four segments of interest are presented in Figure (6.118)

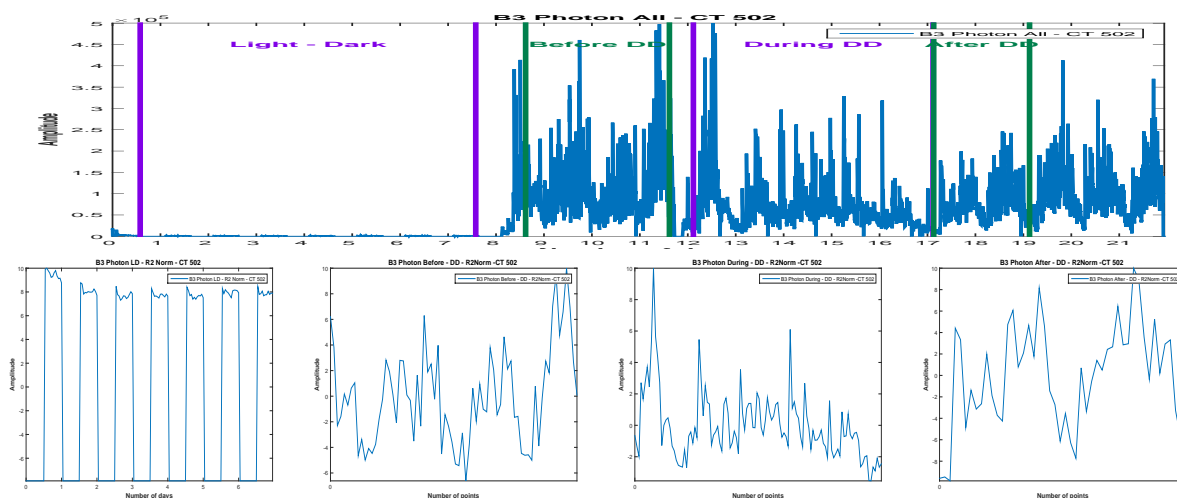


Figure 6.118: Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled

For the LD part, 7 days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.119): Via the FFT method, the principal period is estimated at 24 hours and the only other non-zero peak in the PC amplitude vector corresponds to 8 hours. The dominant period of 24 hours and the corresponding 8 hours harmonic are also the only non-zero amplitudes in the PC amplitude vector corresponding to the Proposed Method.

We consider now 4-days length signals (windows) from the available signal, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.120).

For all the windows, a 24 hours periodicity is confirmed, via the two methods, Fig-

## 6.6.2 - Dominant Period evolution: Photon

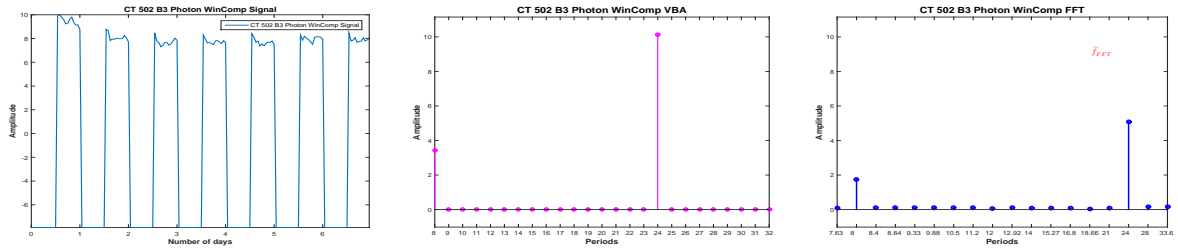


Figure 6.119: Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)

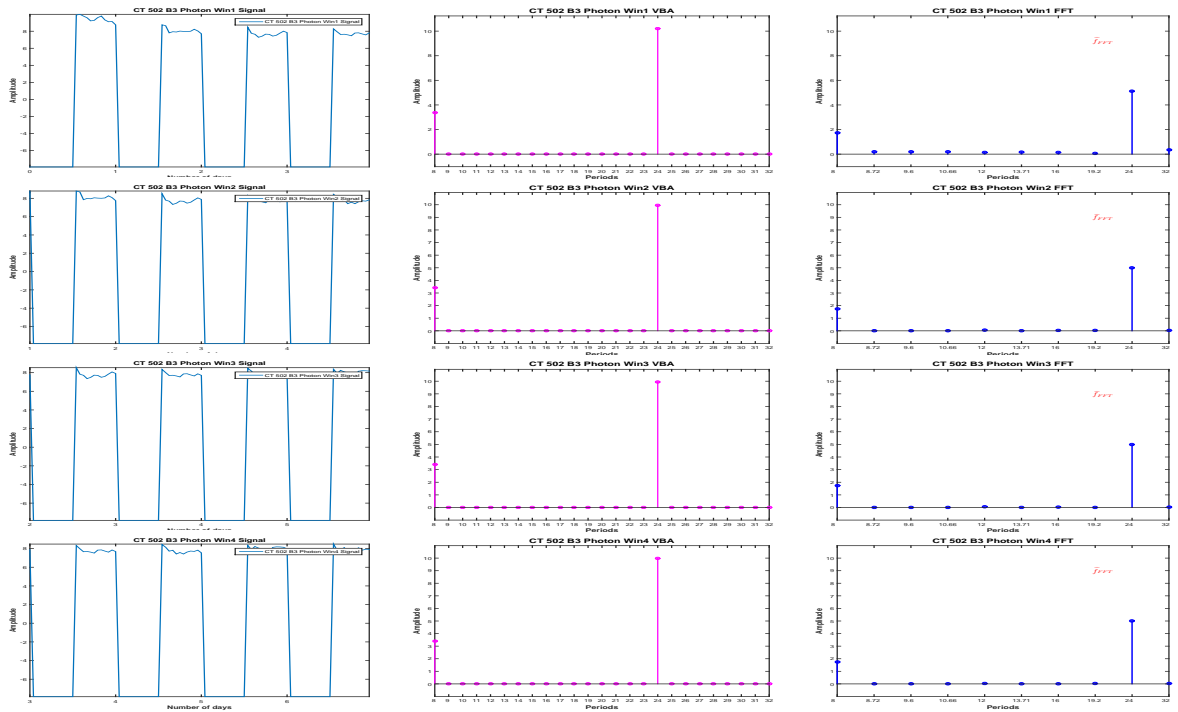


Figure 6.120: PC Stability: PC estimation via FFT and VBA for 4-days length signals

ure (6.121).

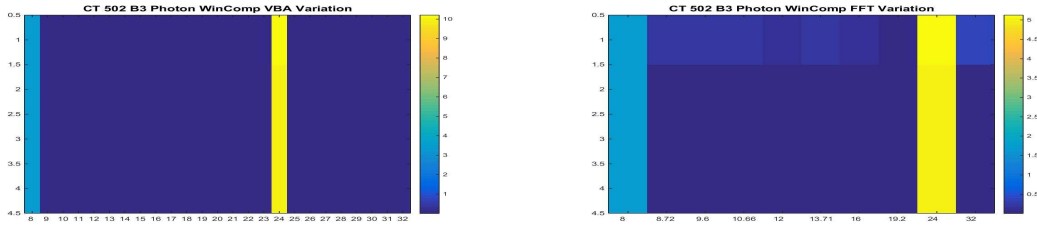


Figure 6.121: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.122):

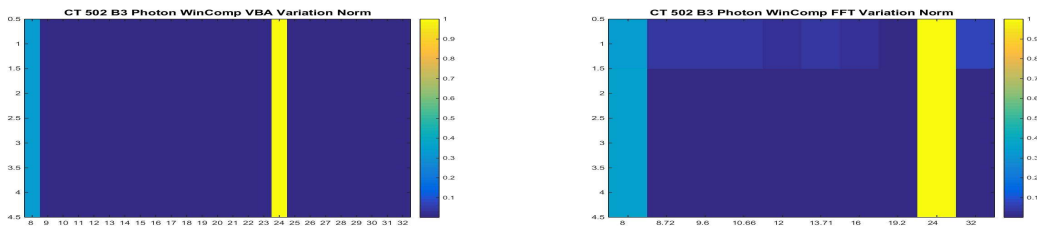


Figure 6.122: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

We consider now the DD period. First, we consider the **before** treatment segment. Only 3 days are available for this segment. We compute the PC vector, using FFT and Proposed Method, Figure (6.123): Via the FFT, the highest pick is set at 24. Other important peaks

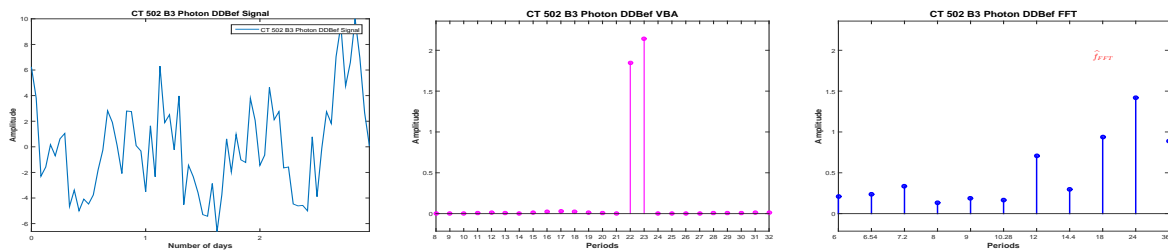


Figure 6.123: DD **before** signal (a) and the corresponding PC via VBA (b) and FFT (c)

correspond to 18, 36 and 12 hours. Of course, 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 23 hours and the other non-zero peak in the PC amplitude vector corresponds to 22 hours.

For the **during** treatment part of the data, five days are available. We compute the PC corresponding to the signal using the Proposed Method and also using the FFT, Figure (6.124): Via FFT, we obtain a 24 hours periodicity, while via the Proposed Method, the dominant peak corresponds to 23. The only other non-zero amplitude in the PC amplitude vector, corresponds to 30 hour, which is consistent with the next highest peak in the FFT estimation, also 30 hours. We note again the difference between the amplitudes corresponding to the two methods and the sparse PC vector returned by the Proposed Method.



## 6.6.2 - Dominant Period evolution: Photon

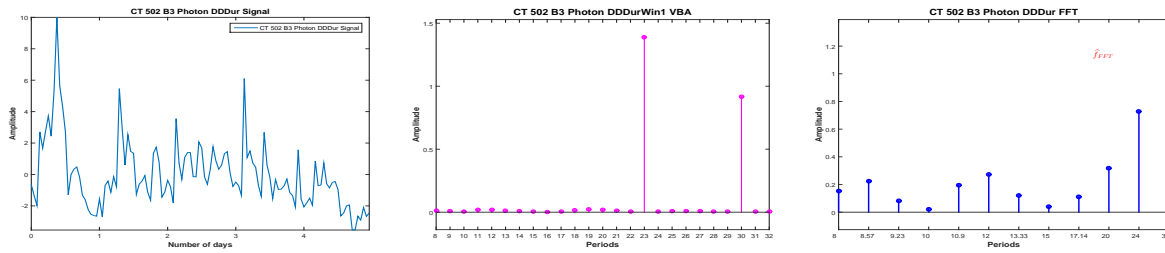


Figure 6.124: DD **during** signal (a) and the corresponding PC via VBA (b) and FFT (c)

We consider now 4-days length signals (windows) from the available **during** treatment part of the data, with a shift of one day and compute the PC via FFT and the Proposed method, Figure (6.125).

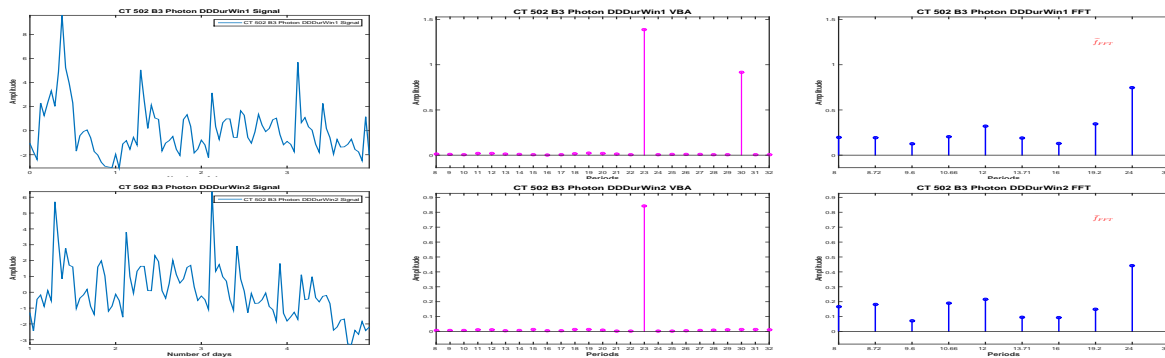


Figure 6.125: PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during

Analysing the **during** segment using the moving window, a 24 hour periodicity is found for both windows, via FFT. However, via the Proposed Method, the periodicity, also stable, is estimated at 23 hours, Figure (6.126)

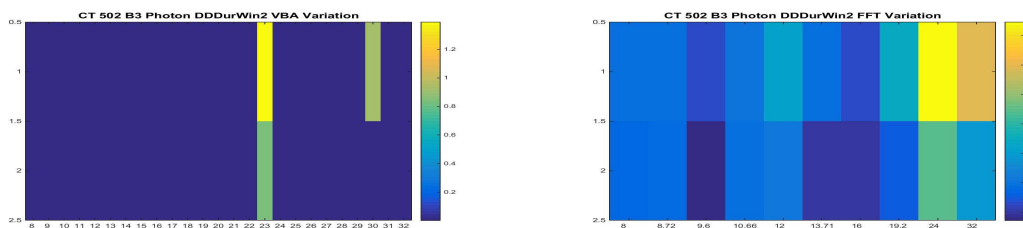


Figure 6.126: PC Stability: Proposed Method (a) vs. FFT (b)

An image of dominant period stability or variability, regardless the numerical value of the amplitude associated with the dominant period is presented in Figure (6.126):

We now consider the **after** treatment part of the data. We have a 2-days length signal. In this case, the dominant period is estimated at 24 hours via FFT. This value is very uncertain, for this particular case (a 2-days length signal) all the interval (16:48) having only one value that is appearing the FFT PC vector, i.e. 24 hours. Via the Proposed Method, the dominant period is set at 22 hours.

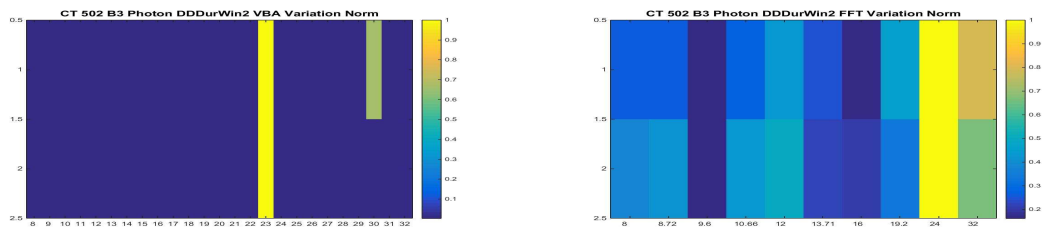


Figure 6.127: PC Stability: Proposed Method (a) vs. FFT (b), normalized results

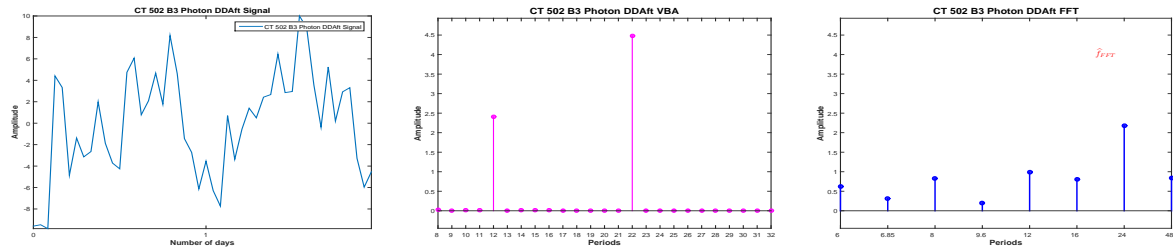


Figure 6.128: DD **after** signal (a) and the corresponding PC via VBA (b) and FFT (c)

The evolution of the dominant period, estimated via the Proposed Method and FFT, during the complete Photon signal, for each segment of interest and for four-days length signals for the segments that allowed this analysis is presented in Table (6.12).

### 6.6.3 Conclusions: Dominant Period evolution

We compared the FFT and Proposed Method PC amplitude vector evolution during the four segments of interest for mouse CT 503 B3, considering the gene expression (Photon Absorption signal) and the rest-activity patterns (Activity signal), Figure (6.106). For both data, we have considered four segments of interest, i.e. the LD part of the signal, corresponding to the synchronization and the DD part of the signal, corresponding to the complete darkness, for which we have considered the segments corresponding to the **before**, **during** and **after** treatment phase, Figure (6.107) (rest-activity patterns) and Figure (6.118).

In particular, for the first segment of interest, i.e. the LD part of data and for the **during** treatment part, we have considered four days length signals to analyse the stability or variability of the dominant period for the considered segment itself.

The dominant period evolution, corresponding to the Proposed Method and FFT for the rest-activity patterns (Activity) signal is presented in Table (6.11). Via FFT, the dominant period seems stable, at 24 hours, during the experiment, except the **after** treatment part of the data, for which the dominant period is estimated at 12 hours. Via the Proposed Method, the period corresponding to the **after** treatment part is also estimated at 12 hours, in accordance with the FFT result. For the other part of the data, the dominant period is also estimated around 24 hours, but with small variations between 23 and 25 hours.

The dominant period evolution, corresponding to the Proposed Method and FFT for the gene expression data (photon absorption signal), is presented in Table (6.12). For this experiment, via FFT, the dominant period is 24 hours for each segment, and seems stable. Via the Proposed Method, slight variations are detected during the experiment: a 23 hour periodicity corresponding to the **before** and **during** treatment part of data and a 22 hour periodicity

### 6.6.3 - Conclusions: Dominant Period evolution

---

CT 502 - B3	Dominant Period - Proposed Method	Dominant Period - FFT
LD complete	24	24
LD win 1	24	24
LD win 2	24	24
LD win 3	24	24
LD win 4	24	24
DD before	23	24
DD during	23	24
DD during win 1	23	24
DD during win 2	23	24
DD after	22	24

Table 6.12: CT 502 B3: Photon Dominant Period Stability: Proposed Method vs. FFT

corresponding to the **after** treatment part of data.

# 7

## Conclusions and perspectives

### Contents

---

<b>7.1 Conclusions</b> . . . . .	<b>179</b>
<b>7.2 Perspectives</b> . . . . .	<b>181</b>
7.2.1 Multicomponent case: PC amplitudes vectors for multiple inputs . . .	181
7.2.2 Sparsity Rate: initialisation and week dependencies . . . . .	182

---

### 7.1 Conclusions

---

THE fundamental objective of this thesis was the study of the periodic components (PC) amplitudes vector corresponding to chronobiological signals. The context of our work was chronobiological signal recorded in cancer treatment experiments, and the domain of interest was the circadian domain. A very important prior information offered by the biologists concerns the PC amplitudes vector: the number of periods modulating the recorded signals is small. Also, the experiments impose two fundamental limitations: the short duration of the signals (relative to the prior dominant period) and the high level of noise. Moreover, for studying the stability or variability of the PC amplitudes vector, even shorter signals have to be considered.

In this context, the objective was to develop a method that can offer a good PC amplitudes vector estimation with a certain precision. During this thesis, one hour precision was considered.

Translating the biological prior information of the small numbers of non zero elements in PC amplitudes vector as sparsity, the problem considered in this thesis has been formulated as it follows:

*Estimate the sparse PC amplitudes vector corresponding to very short duration signals relative to the (prior) dominant period with a given precision. The number of components is small (sparse PC amplitudes vector) but unknown.*

First, we have examined the classical methods and identified their drawbacks in the context of our subject. The Fourier Transform methods are successfully used if the duration of the considered signals are long enough. In our case, due to the short duration of the signals, the precision is far from being satisfactory. In Chapter (3), Section (3.2), we have showed that for a four days length signal, in the interval of interest, i.e. the circadian domain  $[8h : 32h]$ , we obtain informations corresponding only to 10 periods. Zero padding technique was considered in Section (3.3), but in this case, the estimation is not sparse, and selecting periods is difficult. Direct and generalized Inversion was considered in Section (3.4), and we have showed that except the case where the considered signal is not noisy, the method fails to offer a good estimation of the PC amplitudes vector, even though the reconstruction of the signal is accurate for all levels of noise. In Section (3.5) we have considered the Least Squares and the regularized Least Squares, showing that both approaches fail to correctly estimate the PC amplitude vector. In Section (3.6) we considered the  $L_1$  regularization and the LASSO method. This method is the nearest state of art method to account for sparsity through  $L_1$  regularization term. However, we showed that even if very often the results are good, it happens also that for different noise realizations the algorithm does not give satisfactory results.

In the approach proposed in the this thesis, we have considered a forward model, built in the same way as the one corresponding to the Fourier transform, but using the elements that correspond to the considered circadian PC vector, Equation (4.4) and Equation (4.7). Estimating the PC amplitudes vector from the linear model corresponds to an ill-posed inverse problem. More precisely, in the context of this forward model, the duration of the observed signal (96h in synthetic simulations) the limits and the precision of the periods to explore  $[8h : 32h]$ , the matrix of the forward model has a huge condition number. We have proposed an inversion based on general Bayesian Inference, building an hierarchical model based on the generalized Student-t prior distribution modelling the PC amplitudes vector.

The Student-t distribution was considered as a sparsity enforcing prior due to its IGSM property, i.e. the standard Student-t distribution  $St(x|\nu)$  can be expressed as the marginal of a bivariate Normal-Inverse Gamma distribution, when considering equal parameters of the Inverse Gamma distribution  $\alpha = \beta = \frac{\nu}{2}$ . In this thesis, we have used a generalized form of the Student-t distribution, introduced via the IGSM property but not imposing the equality between the shape and scale parameters corresponding to the Inverse Gamma distribution. The advantage was given by the fact that for the generalized Student-t distribution, the variance can take any positive value, (different from the standard Student-t distribution for which the variance is greater than 1). Some details about this generalized Student-t distribution and its properties are presented in Appendix (A) and in a paper submitted to IEEE Signal Processing Letters.

The Bayesian approach allowed us to estimate also the variances associated with the PC amplitude vector. The model that was developed in this thesis considered a non-stationary Gaussian model for the errors (noise).

In Chapter (5) we have presented the results corresponding to the proposed model, showing that the PM estimation corresponding to the proposed Student-t hierarchical model gives accurate results which are stable.

In Chapter (6) the developed algorithm was applied on real data issued from cancer treatment experiments. In this chapter, the results corresponding to the proposed method are compared to the FFT results, which is the standard method for PC estimation in chronobiology today. We have showed that the proposed method is able to detect variances that are not detected by the FFT method. Some details of this approach and its application on simulated and real data are published in the following papers:

- Mircea Dumitru, Ali Mohammad-Djafari, Simona Baghai Sain (2015). *Precise periodic components estimation for chronobiological signals through Bayesian Inference with sparsity enforcing prior*. EURASIP Journal on Bioinformatics and Systems Biology, special issue on Bayesian Methods for Computational Systems Biology, 2015
- Ali Mohammad-Djafari, Mircea Dumitru (2015). *Bayesian sparse solutions to linear inverse problems with non-stationary noise with Student-t priors*, Digital Signal Processing, 2015, doi:10.1016/j.dsp.2015.08.005

## 7.2 Perspectives

In this subsection we discuss some possible perspectives of our work.

### 7.2.1 Multicomponent case: PC amplitudes vectors for multiple inputs

We discussed during this thesis, in Chapter (4), the fact the two equivalent direct models can be considered: the one expressed in Equation (4.4) and the one expressed in Equation (4.7), equivalent with Equation (4.4) via (4.6). The same hierarchical model, with a generalized Student-t distribution as a sparsity enforcing prior for the PC amplitude vector and a non-stationary Gaussian model for the errors can be applied for both forward models. However, in the case of the linear model Equation (4.4), the link between the two PC amplitudes vectors (i.e.  $\mathbf{f}_1$  and  $\mathbf{f}_2$ , corresponding to the sine and cosine) is done by imposing the same variance  $\mathbf{v}_f$ . This idea can be applied when the developed algorithm is used not for the case of one signal recorded in the experiment, but for two or more signals. For the same experiment, the gene expression signal and the activity signal are available. For the same experiment other signal associated with the circadian clock can be recorded, (like temperature). The developed algorithm can be extended to the multicomponent case. Given a set of available signals, recorded for the same experiment, the algorithm will estimate the PC amplitudes vectors containing the most common periodic components corresponding to the set of signals. This can be done via the same technique used in the case of one available signal, imposing the same variance for every PC amplitude vector corresponding to the available signals.

In the following, for the sake of simplicity we will consider as a reference, the linear model expressed in Equation (4.7). However, the same development can be done using Equation (4.4). If the number of available signals is  $K$ , and for each available signal we use the notation  $\mathbf{g}_k, k \in \{1, 2, \dots, K\}$ , for the corresponding PC amplitudes vector we use the notation  $\mathbf{f}_k, k \in \{1, 2, \dots, K\}$  and for the corresponding error vector  $\mathbf{\epsilon}_k, k \in \{1, 2, \dots, K\}$  the developed method can be used in order to estimate each  $\mathbf{f}_k$ , for  $k \in \{1, 2, \dots, K\}$ , considering as the

## 7.2.2 - Sparsity Rate: initialisation and week dependencies

---

direct model  $\mathbf{g}_k = \mathbf{H}\mathbf{f}_k + \boldsymbol{\epsilon}_k$ , for  $k \in \{1, 2, \dots, K\}$  and the corresponding hierarchical model. However, doing this, the advantage of multiple signals available is not exploited, since for estimation  $\mathbf{f}_k$  it is used only the corresponding signal  $\mathbf{g}_k$ .

In order to use all the available data, we can link all the PC amplitudes vectors estimations  $\mathbf{f}_k$ , for  $k \in \{1, 2, \dots, K\}$  by imposing the same variance:

$$\mathbf{v}_{\mathbf{f}_k} = \mathbf{v}_{\mathbf{f}}, \forall k \in \{1, 2, \dots, K\}; \quad (7.1)$$

The advantage of such an approach is given directly by the sparsity mechanism, explained in Section (4.1): during the iterations the algorithm will impose zero value amplitudes for the periods in the PC amplitude vector corresponding to small variances. Therefore, the final estimation for each PC amplitudes vector will contain non-zero values only for the periods that correspond to the most signals and will exclude the ones corresponding to a small number of signals from the available data.

Formally, this can be done using as the forward model:

$$\mathbf{g}_k = \mathbf{H}\mathbf{f}_k + \boldsymbol{\epsilon}_k, k \in \{1, 2, \dots, K\}; \quad (7.2)$$

or

$$\mathbf{G} = \mathbf{H}\mathbf{F} + \boldsymbol{\Psi} \quad (7.3)$$

where  $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_K]$ ,  $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_K]$  and  $\boldsymbol{\Psi} = [\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \dots, \boldsymbol{\epsilon}_K]$ .

## 7.2.2 Sparsity Rate: initialisation and week dependencies

Another perspective of the our work related with the generalized Student-t distribution. The  $St_g$  distribution was used as the prior distribution in our hierarchical model in order to enforce sparsity. The distribution is presented in detail in Appendices (A).

Briefly, we have noted that for the standard Student-t distribution  $St(x|\nu)$  its variance is defined as  $\frac{\nu}{\nu-2}$  for  $\nu > 2$  and  $\infty$  for  $1 < \nu \leq 2$ , so the variance of the standard Student-t distribution is always superior to 1. We have also noted that for the generalized Student-t distribution  $St_g(x|\alpha, \beta)$ , presented in Equation (A.25), its variance is defined as  $\frac{\beta}{\alpha-1}$  for  $\alpha > 1$  and  $\infty$  for  $1/2 < \alpha \leq 1$ , (and in particular has the same expression as the mean of the Inverse Gamma distribution corresponding to the Normal-Inverse Gamma distribution that is marginalized for obtaining the generalized Student-t distribution) so once the shape parameters  $\alpha$  is set greater than 1, the variance of  $St_g(x|\alpha, \beta)$  can be set at any wanted positive value. In particular, in the context of a distribution that is used as a sparsity enforcing prior this is of great interest.

For the developed algorithm, a crucial interest is represented by the initialization part, where the prior parameters are set. Via the generalized Student-t distribution one way to consider the initialization is to set the two parameters via the mean and the variance of the Inverse Gamma. Noting that the mean of the Inverse Gamma distribution has the same expression as the variance of the Student-t distribution:

$$E_{\mathcal{IG}}[v] = \frac{\beta}{\alpha-1} = Var_{St_g}[x] = \epsilon, \text{ for } \alpha > 1 \quad (7.4)$$

and the expression of the Inverse Gamma variance is

$$Var_{\mathcal{IG}}[v] = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} = w, \text{ for } \alpha > 2, \quad (7.5)$$

the choice of the prior parameters can be done in accordance with data, initializing the algorithm via the associated  $St_g$  variance of the unknown vector, i.e. the PC amplitudes vector  $\mathbf{f}$  and the  $IG$  variance of the corresponding variance,  $\mathbf{v}_f$ .

In our work, the two prior parameters corresponding to the generalized Student-t distribution are set via  $\epsilon$ , which denotes the Student-t variance and  $w$ , which denotes the Inverse Gamma variance. We have used Equation (4.65), in such way that  $\alpha > 2$ , so that both the variances are well defined and the Student-t variance is small (in order to assure the sparsity, Figure (A.3), Appendices (A)). One possible extension is to use Equation (4.65) with variance values  $\epsilon$  and  $w$  statistically derived from data.

Expressing the two parameters of the distribution using the mean and the variance of the corresponding Inverse Gamma distribution, we have showed that the Sparsity Rate (SR) depends on  $\epsilon$ , the mean of the Inverse Gamma and the variance of the generalized Student-t distribution (Figure (A.3), Appendices (A)). But Figure (A.3) shows another interesting result, a weak dependence of the SR with respect to  $w$ , the variance of the Inverse Gamma distribution. This result should be deepened, formalized and analytically explained by considering the expression  $SR = p(|x| > t)$ .





# A

## Generalized Student-t

### Contents

---

<b>A.1 Context</b> . . . . .	<b>185</b>
<b>A.2 Student-t: Origins</b> . . . . .	<b>187</b>
A.2.1 ML parameter estimation . . . . .	187
A.2.2 Bayesian parameter estimation . . . . .	188
<b>A.3 Generalization from the Bayesian point of view</b> . . . . .	<b>189</b>
<b>A.4 Sparsity Rate and Generalized Student-t</b> . . . . .	<b>190</b>
<b>A.5 Conclusion</b> . . . . .	<b>192</b>

---

THIS Appendix is dedicated to the generalized Student-t distribution. During this thesis, we have used the notation  $\mathcal{S}t_g(x|\alpha, \beta)$  for the generalized Student-t, which was obtained from the bivariate Normal-Inverse Gamma distribution. The interest for this generalization is the possibility to set any positive value for the variance, so therefore imposing the wanted sparsity rate. In this Appendix, first we discuss the context where this distribution can be used, then we discuss the origins of the Student-t distribution, then we present the computations for obtaining the analytical expression of the generalized Student-t distribution and the computations corresponding to the expressions of the mean and variance.

### A.1 Context

---

During this thesis we have noted that beside the regularization techniques, the other approach that has been successfully tested to account for sparsity is the Bayesian one. We have also mentioned that for Bayesian inference generally two strategies are used: **(a)** the strict sparsity requirement, via Bernouilli distribution: Bernouilli-Gaussian [GDI90, CGI96, GILC08], Bernouilli-Laplace [DB08], Bernouilli-Gamma and **(b)** the sparsity enforcing, where mainly

three categories of priors have been considered and used: Generalized Gaussian (GG), Mixture models and heavy tailed probability laws such as Student-t. A review of those priors can be found in [Tip01, MD12].

The context of the generalized Student-t is placed relative to this second strategy, where the searched solution is not strictly sparse, but sparse in the sense that the a high number of elements are sufficiently small. We note that in this case we, the threshold is particular important, in order to control the interval of small values that are associated with the zero values.

Also, depending on the application considered, the input  $\mathbf{f}$  may have different sparsity rates, depending on the application, therefore the prior information concerning the Sparsity Rate (SR) have to be accounted.

In this Appendices, we present the generalization of the Student-t distribution that allows accounting for the SR. The standard Student-t pdf is given by the expression:

$$\mathcal{S}t(x|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad \nu > 0, \quad (\text{A.1})$$

where  $\nu$  is the number of degrees of freedom and  $\Gamma$  is the Gamma function. It is generalized by introducing a location parameter  $\mu$  and a scale parameter  $\sigma$ :

$$\mathcal{S}t(x|\nu, \mu, \sigma) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \sigma \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{1}{\nu} \left(\frac{x-\mu}{\sigma}\right)^2\right)^{-\frac{\nu+1}{2}}, \quad (\text{A.2})$$

$\nu > 0, \sigma > 0, \mu \in R$

The Student-t distribution can also be expressed via the IGSM:

$$\mathcal{S}t(x|\nu) = \int_0^\infty \mathcal{N}(x|0, \frac{1}{z}) \mathcal{G}(z|\frac{\nu}{2}, \frac{\nu}{2}) dz \quad (\text{A.3})$$

where  $\mathcal{N}(x|m, v) = (2\pi v)^{-\frac{1}{2}} \exp\{-\frac{1}{2v}(x - m)^2\}$  is the Normal distribution and  $\mathcal{G}(x|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}$  is the Gamma distribution. The IGSM name comes from the fact that when a bivariate pdf composed of the product of a zero mean Normal distribution  $\mathcal{N}(x|0, \frac{1}{z})$  with variance  $v = \frac{1}{z}$  and a Gamma pdf for the inverse of the variance (precision)  $z = \frac{1}{v}$ , having equal scale and shape parameters  $\alpha = \beta = \frac{\nu}{2}$

$$p(x, z|\nu) = \mathcal{N}(x|0, \frac{1}{z}) \mathcal{G}(z|\frac{\nu}{2}, \frac{\nu}{2}). \quad (\text{A.4})$$

Now, marginalizing  $z$  or  $v = \frac{1}{z}$  results to Student-t pdf as is given in Equation (A.3).

There are many works on the origins, properties and many other relations and generalizations of this distribution. Some of them can be seen in [Fis25],[Li57],[Zel76],[Fer99],[KBJ00],[AH06]. Those works can mainly be classified in three categories: Probability and information theory, Classical statistics and Bayesian inference. In probability theory Student-t pdf is defined as the pdf of the ratio of two variables:  $z$  with a Normal pdf and  $v$  with a chi-squared pdf. In classical statistics, its origins goes back to the Maximum Likelihood (ML) estimate of the mean and variance of a Normal distribution. In Bayesian inference framework, the origin goes back to the same context of ML parameter estimation but with a Jeffreys prior. In this paper, we

focus on the Bayesian inference of the mean and the variance of a Normal distribution when a Normal-Inverse Gamma prior is assigned to them. By doing this, we obtain an expression for the marginal posterior pdf of the mean which extends the classical definition of the Student-t. In this way, we obtain a two parameters Student-t that we called generalized Student-t  $St_g$ . The rest of Appendices (A) is organized as follows: in Section (A.2), we give a brief presentation of the origins of the distribution considered. Then, in Section (A.3) we present the new generalized Student-t and some of its properties, in particular the computation of the expected value and the variance. We will show that the variance can take any positive values. We also present the link between the variance corresponding to the  $St_g$  and the expected value corresponding to the  $\mathcal{IG}$ , modeling the prior variance. Section (A.4) presents how using this property of the variance can be linked with the SR and presents the relation between those two values. The conclusions are presented in Section (A.5).

---

## A.2 Student-t: Origins

---

Student-t distribution was first derived as a Bayesian posterior distribution in 1876 by Helmert [Hel75, Hel76b, Hel76a] and Lüroth.[L76]. It takes its name from William Sealy Gosset's 1908 paper in Biometrika under the pseudonym "Student" [Gos08, Mor08]. In Probability theory, Student's t-distribution with  $\nu$  degrees of freedom is defined as the distribution of a random variable  $X$  defined as the ratio of two independent random variables:

$$X = \frac{Z}{\sqrt{V/\nu}} = Z\sqrt{\frac{\nu}{V}}, \quad (\text{A.5})$$

where  $Z$  has a standard normal distribution with expected value 0 and variance 1;  $V$  has a chi-squared distribution with  $\nu$  degrees of freedom [SR94]. Student's t-distribution has also been defined as the Maximum Entropy probability distribution for a variable  $X$  for which  $E\{\ln(\nu + X^2)\}$  is fixed [Haz01, VR02, GCSR03, WMMY02]. Also, it has been shown that the T-distribution pdf is a solution to the following differential equation [HMC02]:

$$\begin{cases} (\nu + x^2) f'(x) + (\nu + 1)xf(x) = 0, \\ f(1) = \frac{\nu^{\nu/2}(\nu+1)^{-\frac{\nu}{2}-\frac{1}{2}}}{B(\frac{\nu}{2}, \frac{1}{2})} \end{cases} \quad (\text{A.6})$$

In this paper, we consider two main origins of this distribution: the classical Maximum Likelihood (ML) parameter estimation of a Normal distribution and the Bayesian approach of this problem via the Jeffreys prior. We introduce these two approaches briefly.

### A.2.1 ML parameter estimation

If  $x_1, \dots, x_N$  are i.i.d samples from a Normal distribution with unknown mean  $\mu$  and variance  $v$ ,

$$x_i \sim \mathcal{N}(x_i|\mu, v), \quad i \in \{1, 2, \dots, N\}, \quad (\text{A.7})$$

the maximum likelihood estimate of  $(\hat{\mu}, \hat{v})$  is defined as

$$(\hat{\mu}, \hat{v})_{ML} = \arg \max_{(\mu, v)} \{p(\mathbf{x}|\mu\mathbf{1}, v\mathbf{I})\} = \arg \min_{(\mu, v)} \{\mathcal{L}(\mu, v)\} \quad (\text{A.8})$$

## A.2.2 - Bayesian parameter estimation

---

where  $\mathcal{L}(\mu, v) = -\ln p(\mathbf{x}|\mu\mathbf{1}, v\mathbf{I}) = -\sum_{i=1}^N \ln(p(x_i|\mu, v))$ . It is easy to show that the expression of  $\mathcal{L}(\mu, v)$  depends only on the two following sufficient statistics:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{and} \quad s^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (\text{A.9})$$

and are given by:

$$\begin{cases} \hat{\mu} = \bar{x} \\ \hat{v} = s^2 \end{cases} \quad (\text{A.10})$$

It can also be easily shown that

$$\begin{cases} \bar{x} \sim \mathcal{N}(\bar{x}|\mu, \frac{v}{N}) \\ s^2 \sim \chi^2(s^2|N-1), \end{cases} \quad (\text{A.11})$$

where  $\chi^2$  represents the Chi-squared distribution which is a particular case of the Gamma distribution, and that the random variable defined by  $T = \frac{\bar{x} - \mu}{s/\sqrt{N}}$ , called pivot variable, follows a Student-t distribution with  $N - 1$  degrees of freedom:

$$T = \frac{\bar{x} - \mu}{s/\sqrt{N}} \sim \mathcal{St}(t|N-1). \quad (\text{A.12})$$

## A.2.2 Bayesian parameter estimation

The Bayesian way to estimate  $(\mu, v)$  needs the expression of the likelihood as before:

$$p(\mathbf{x}|\mu, v) = \prod_{i=1}^N \mathcal{N}(x_i|\mu, v) = \mathcal{N}(\mathbf{x}|\mu\mathbf{1}, v\mathbf{I}), \quad (\text{A.13})$$

but also assigning a prior to  $(\mu, v)$ . When choosing the Jeffreys priors  $p(\mu, v) \propto \frac{1}{v}$ , which is a particular case of a Normal-Inverse Gamma pdf:  $p(\mu, v) = \mathcal{N}(\mu|\mu_0, v)\mathcal{IG}(v|\alpha, \beta)$  with  $\mu_0 = \alpha = \beta = 0$ , corresponding to a zero mean Normal distribution and a degenerate case of the Inverse Gamma distribution, i.e.

$$\begin{cases} p(\mu|v) = \mathcal{N}(\mu|0, v) \\ p(v) = \mathcal{IG}(v|0, 0) \end{cases} \quad (\text{A.14})$$

we obtain the expression of the posterior law:

$$p(\mu, v|\mathbf{x}) \propto v^{-\frac{N}{2}-1} \exp\left\{-\frac{1}{2v} \sum_{i=1}^N (x_i - \mu)^2\right\} \quad (\text{A.15})$$

from which we can deduce:

$$\begin{cases} p(\mu|\mathbf{x}, v) = \mathcal{N}(\mu|\bar{x}, \frac{v}{N}) \\ p(v|\mathbf{x}) = \chi^2(v|N) \end{cases} \quad (\text{A.16})$$

Now, looking for the marginal:

$$\begin{aligned} p(\mu|\mathbf{x}) &= \int_0^\infty p(\mu, v|\mathbf{x}) dv \propto \int_0^\infty \mathcal{N}(\mathbf{x}|\mu\mathbf{1}, v\mathbf{I}) \frac{1}{v} dv \\ &= \mathcal{St}(\mu|\bar{x}, \sqrt{Ns^2}, N). \end{aligned} \quad (\text{A.17})$$

The standard Student-t appears via the pivot variable,  $T$ :

$$T = \frac{\bar{x} - \mu}{\sqrt{s^2}/\sqrt{N}} \sim \mathcal{St}(t|N-1). \quad (\text{A.18})$$

## A.3 Generalization from the Bayesian point of view

---

In the previous Bayesian estimation, a more general informative prior for  $(\mu, v)$  in the form

$$p(\mu, v) = p(\mu|v)p(v) \quad (\text{A.19})$$

with

$$\begin{cases} p(\mu|v) = \mathcal{N}(\mu|0, v) \\ p(v|\alpha, \beta) = \mathcal{IG}(v|\alpha, \beta) \end{cases} \quad (\text{A.20})$$

leads to the joint posterior distribution:

$$\begin{aligned} p(\mu, v|\mathbf{x}) &\propto \prod_{i=1}^N \mathcal{N}(x_i|\mu, v) \mathcal{N}(\mu|0, v) \mathcal{IG}(v|\alpha, \beta) \\ &\propto \mathcal{N}(\mu|\hat{\mu}, \hat{v}) \mathcal{IG}(v|\hat{\alpha}, \hat{\beta}) \end{aligned} \quad (\text{A.21})$$

with

$$\begin{cases} \hat{\mu} = \frac{N}{N+1} \bar{x} \\ \hat{v} = \frac{1}{N+1} v \end{cases} ; \begin{cases} \hat{\alpha} = \alpha + \frac{N}{2} \\ \hat{\beta} = \beta + \frac{1}{2} \left( -\frac{N^2}{N+1} \bar{x}^2 + \sum_{i=1}^N x_i^2 \right) \end{cases} \quad (\text{A.22})$$

From this joint posterior distribution the expression of the conditional posterior  $p(\mu|\mathbf{x}, v)$  and posterior  $p(v|\mathbf{x}, \alpha, \beta)$  are easily derived:

$$\begin{cases} p(\mu|\mathbf{x}, v) = \mathcal{N}(\mu|\hat{\mu}, \hat{v}) \\ p(v|\mathbf{x}, \alpha, \beta) = \mathcal{IG}(v|\hat{\alpha}, \hat{\beta}) \end{cases} \quad (\text{A.23})$$

The marginal posterior distribution  $p(v|\mathbf{x}, \alpha, \beta)$  is an Inverse Gamma distribution and is described by the two parameters  $\hat{\alpha}$  and  $\hat{\beta}$  of Equation (A.22). For obtaining the marginal posterior distribution corresponding to  $\mu$ , we need to compute:

$$p(\mu|\mathbf{x}) = \int_0^\infty \mathcal{N}(\mu|\hat{\mu}, \hat{v}) \mathcal{IG}(v|\hat{\alpha}, \hat{\beta}) \, dv \quad (\text{A.24})$$

This is how a generalization of the Student-t appears naturally. In the following we consider Equation (A.24) for which we show the analytical expression and some of its properties. For the sake of simplicity we develop the case corresponding to  $\hat{\mu} = 0$  corresponding to the zero mean Normal distribution.

**Theorem:** For two random variables following a bivariate Normal-Inverse Gamma distribution,  $(X, V) \sim \mathcal{N}(x|0, v) \mathcal{IG}(v|\alpha, \beta)$ , the marginal distribution corresponding to  $X$  is a Student-t like distribution, generalizing the classical form, denoted  $\mathcal{St}_g$ , depending on the two parameters  $\alpha$  and  $\beta$  corresponding to the Inverse-Gamma distribution having the following pdf:

$$\begin{aligned} \mathcal{St}_g(x|\alpha, \beta) &= \int_0^\infty \mathcal{N}(x|0, v) \mathcal{IG}(v|\alpha, \beta) \, dv \\ &= \frac{\Gamma(\alpha + \frac{1}{2})}{\sqrt{2\beta\pi} \Gamma(\alpha)} \left( 1 + \frac{x^2}{2\beta} \right)^{-(\alpha + \frac{1}{2})}, \alpha > 0, \beta > 0. \end{aligned} \quad (\text{A.25})$$

## A.2.2 - Bayesian parameter estimation

---

**Proof:** The expression of the distribution is obtained directly:

$$\begin{aligned}
 \mathcal{S}t_g(x|\alpha, \beta) &= \\
 &= \int_0^\infty \frac{1}{\sqrt{2\pi v}} \exp\left\{-\frac{1}{2v}x^2\right\} \frac{\beta^\alpha}{\Gamma(\alpha)} v^{-(\alpha+1)} \exp\left\{-\frac{\beta}{v}\right\} dv \\
 &= \frac{\beta^\alpha \Gamma(\alpha + \frac{1}{2})}{\sqrt{2\pi} \Gamma(\alpha) (\beta + \frac{x^2}{2})^{(\alpha + \frac{1}{2})}} \int_0^\infty \mathcal{IG}\left(v|\alpha + \frac{1}{2}, \beta + \frac{x^2}{2}\right) dv \\
 &= \frac{\Gamma(\alpha + \frac{1}{2})}{\sqrt{2\beta\pi} \Gamma(\alpha)} \left(1 + \frac{x^2}{2\beta}\right)^{-(\alpha + \frac{1}{2})}.
 \end{aligned} \tag{A.26}$$

**Remark 1:** For  $\alpha = \beta = \frac{\nu}{2}$ , we obtain the standard case:  $\mathcal{S}t_g(x|\frac{\nu}{2}, \frac{\nu}{2}) = \mathcal{S}t(x|\nu)$ .

**Remark 2:** The expected value of the  $\mathcal{S}t_g(x|\alpha, \beta)$  is zero for  $\alpha > \frac{1}{2}$  and is undefined otherwise:

$$\mathbb{E}\{x|\alpha, \beta\} = \begin{cases} 0 & \text{for } \alpha > \frac{1}{2}, \\ \text{undefined} & \text{for } 0 < \alpha \leq \frac{1}{2}. \end{cases} \tag{A.27}$$

**Proof:** The proof follows the same steps as the one corresponding to the standard case, also using the fact that  $\mathcal{S}t_g(x|\alpha, \beta)$  is an even function. One difference appears for the interval of the  $\alpha$  parameter for which the mean is defined. In this case, during the computations the integral  $\int_0^\infty x \left(1 + \frac{x^2}{2\beta}\right)^{-(\alpha + \frac{1}{2})} dx$  is considered. The integral is finite only for  $\alpha > \frac{1}{2}$ .

**Remark 3:** The variance of the  $\mathcal{S}t_g(x|\alpha, \beta)$  is  $\frac{\beta}{\alpha-1}$  for  $\alpha > 1$ , is  $\infty$  for  $\frac{1}{2} < \alpha \leq 1$  and is undefined otherwise:

$$\text{Var}\{x|\alpha, \beta\} = \begin{cases} \frac{\beta}{\alpha-1} & \text{for } \alpha > 1, \\ \infty & \text{for } \frac{1}{2} < \alpha \leq 1, \\ \text{undefined} & \text{for } 0 < \alpha \leq \frac{1}{2}. \end{cases} \tag{A.28}$$

**Proof:** The proof follows the same steps as the one corresponding to the standard case and is using the Euler integral of the first kind.

**Remark 4:** We note that for  $\alpha > 1$ , the mean of the  $\mathcal{IG}$  distribution appearing in the expression of the bivariate Normal-Inverse Gamma distribution is well defined and has the same value as the variance of  $\mathcal{S}t_g$ :

$$\text{Var}_{\mathcal{S}t_g}\{x|\alpha, \beta\} = \mathbb{E}_{\mathcal{IG}}\{v|\alpha, \beta\} = \frac{\beta}{\alpha - 1} \tag{A.29}$$

This property is particularly interesting in the sparsity context. It is showing how the sparsity mechanism is strongly related with the associated prior variances of the model. The bivariate Normal-Inverse Gamma distribution and the corresponding marginal, i.e. the generalized Student-t distribution are presented in Figure (A.1).

## A.4 Sparsity Rate and Generalized Student-t

---

The Student-t distribution represents a good prior distribution for enforcing the sparsity, thanks to its heavy tailed form. Among many other distributions having a heavy tailed form,

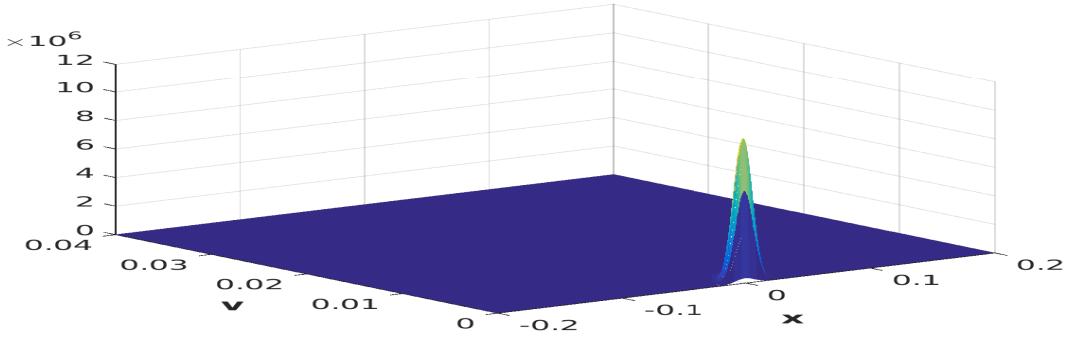
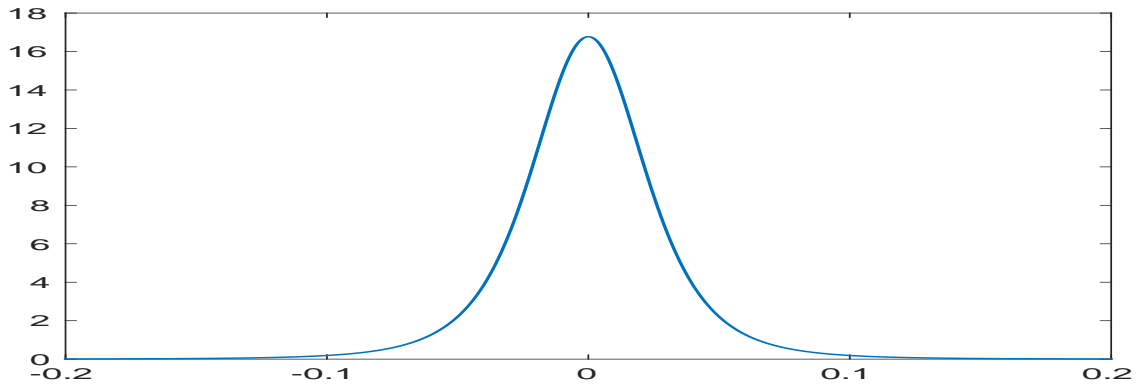

 a) Normal-Inverse Gamma modelling  $(x, v)$ 

 b)  $\mathcal{S}t_g$  modelling  $x$ :  $\text{Var}_{\mathcal{S}t_g}\{(x|\alpha, \beta)\} = \epsilon = 0.1$ 

 Figure A.1: Normal-Inverse Gamma distribution (a) and the corresponding marginal  $\mathcal{S}t_g$  (b). Illustrated for  $\epsilon = 0.1$  and  $k = 1$ , Eq. (A.30a) and Eq. (A.30b)

such as Double Exponential (DE) or Generalized Gaussian (GG), the Student-t distribution is particularly interesting thanks to its IGSM expression and the conjugacy property of the Normal and Gamma (or Inverse Gamma) distributions. Therefore, in many Bayesian approaches searching for sparse solutions, this distribution is of great interest. However, the mechanism of the Bayesian algorithms where a sparse solution is searched is using at one hand the heavy-tailed form of the prior distribution and at the other hand the small value of the variance. The use of the standard Student-t distribution imposes a limitation: the variance is always greater than 1. In this context, considering the generalization of the Student-t distribution and notably the relation from Equation (A.28), we see the importance of this generalization for which the variance may be set at any required value. Furthermore, Equation (A.29) sets the link between the sparsity of a quantity  $x$  and the associated variances  $v$ . For  $\mathcal{S}t_g$  variance  $\epsilon$  and  $\mathcal{IG}$  variance  $k$ , we have:

$$\text{Var}_{\mathcal{S}t_g}\{(x|\alpha, \beta)\} = \text{E}_{\mathcal{IG}}\{(v|\alpha, \beta)\} = \frac{\beta}{\alpha - 1} = \epsilon; \quad (\text{A.30a})$$

$$\text{Var}_{\mathcal{IG}}\{(v|\alpha, \beta)\} = \frac{\beta^2}{(\alpha - 1)^2(\alpha - 2)} = k; \quad (\text{A.30b})$$

A solution for setting the scale and shape parameters  $\alpha$  and  $\beta$  depending on the variance of the generalized Student-t distribution  $\epsilon$  and the variance of the Inverse Gamma distribution  $k$



## A.2.2 - Bayesian parameter estimation

---

is given in Equation (A.31).

$$\begin{cases} \alpha = 2 + \epsilon_0^2 \\ \beta = \sqrt{k}\epsilon_0(1 + \epsilon_0^2) \end{cases}, \epsilon_0 = \frac{\epsilon}{\sqrt{k}}; \quad (\text{A.31})$$

We note that via Equation (A.31),  $\alpha > 2$  and  $\beta > 0$ : for both the  $\mathcal{S}t_g$  and  $\mathcal{IG}$  distributions, the corresponding expected values and variances are well defined. The Sparsity Rate (SR) is defined in Equation (A.32).

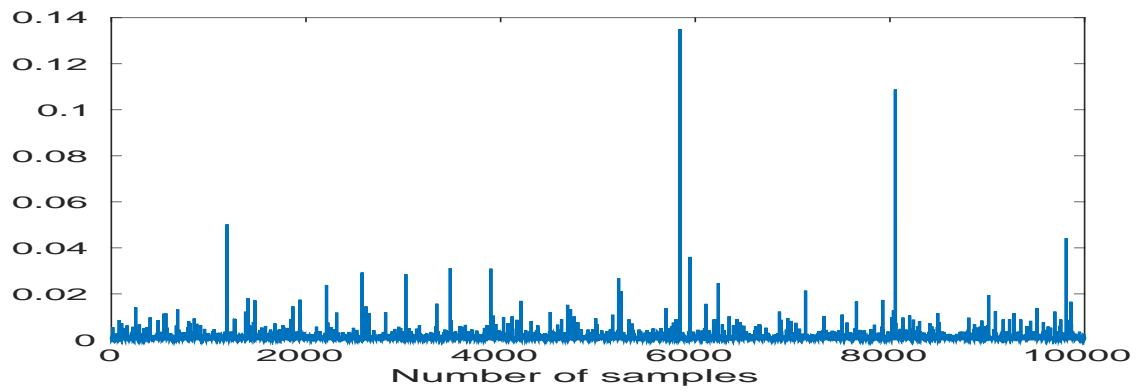
$$\text{SR} = \frac{\#\text{non-zero points}}{\#\text{points}} \quad (\text{A.32})$$

To show the link between the SR and the variance of the generalized Student-t distribution,  $\epsilon$ , we consider the sparse structure  $x$  and the corresponding variance  $v$  as it follows: (i) we consider 10000 samples for the variance,  $\mathcal{IG}$  distributed, Figure (A.2) (a); (ii) we consider 10000 samples for  $x$ ,  $\mathcal{N}$  distributed, with zero mean and the variance selected from (a), Figure (A.2) (b); (iii) we set a threshold for  $x$ , ( $|x| < 0.1$ ), obtaining a sparse structure, Figure (A.2) (c); We consider different sparse structures  $x$ , corresponding to different values  $\epsilon$  and  $k$ , Equation (A.31), and compute the associated SR value, Equation (A.32). Figure (A.3) presents the behaviour of the SR depending on the  $\mathcal{S}t_g$  variance  $\epsilon$ , for different values considered for the variance of the  $\mathcal{IG}$ ,  $k$ . We note that regardless the considered values  $k$ , the SR is decreasing with  $\epsilon$ .

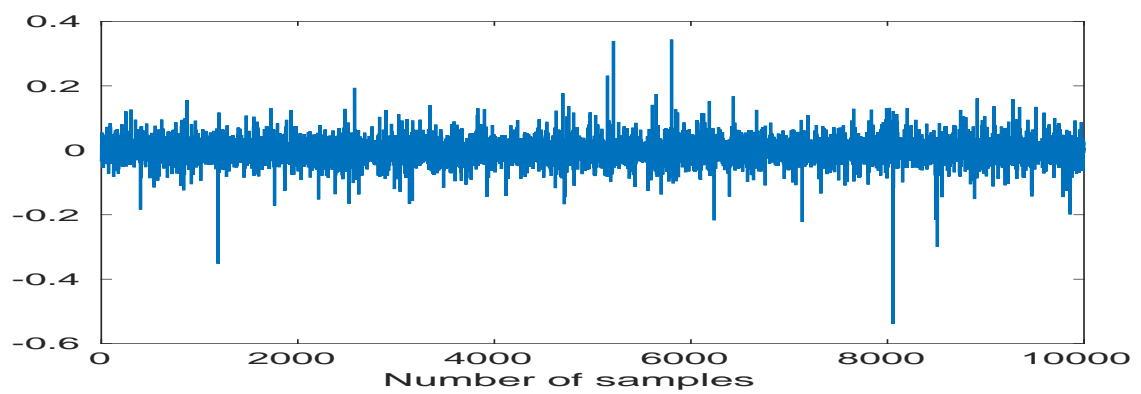
## A.5 Conclusion

---

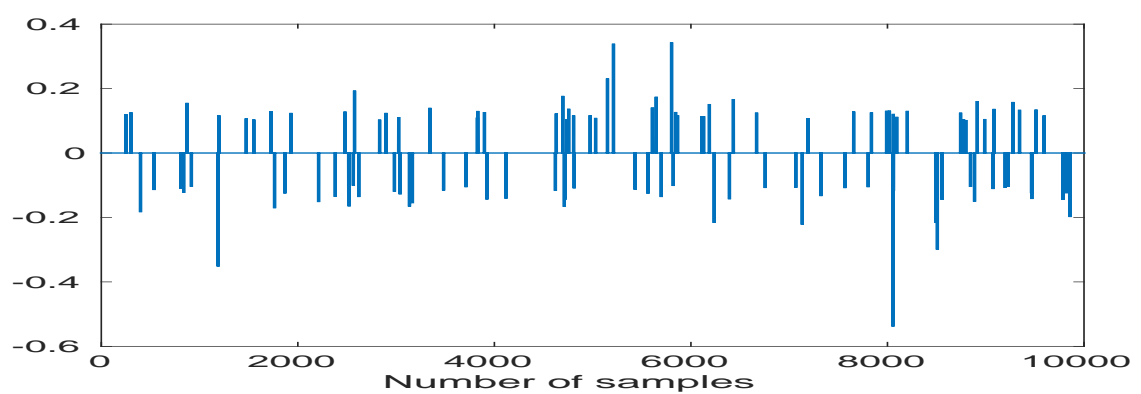
Based on Bayesian inference of the mean  $\mu$  of a Gaussian pdf  $\mathcal{N}(x|\mu, v)$  with unknown variance  $v$  on which an a priori Inverse Gamma pdf  $\mathcal{IG}(v|\alpha, \beta)$  is assigned, we considered the generalized Student-t pdf  $\mathcal{S}t_g(x|\alpha, \beta)$  which becomes the standard Student t-distribution when  $\alpha = \beta = \nu/2$ . We showed that the variance of this distribution can take any positive values, and how it is linked with the SR. We showed how the two parameters of the generalized Student-t can be set such that the value of the variance is imposed. Depending on applications, the SR can have different values, and the proposed distribution is offering a solution for selecting a prior which is in accordance with data in terms of SR.



a)  $\mathcal{IG}$  distributed samples  $v \sim \mathcal{IG}(v|\alpha, \beta)$



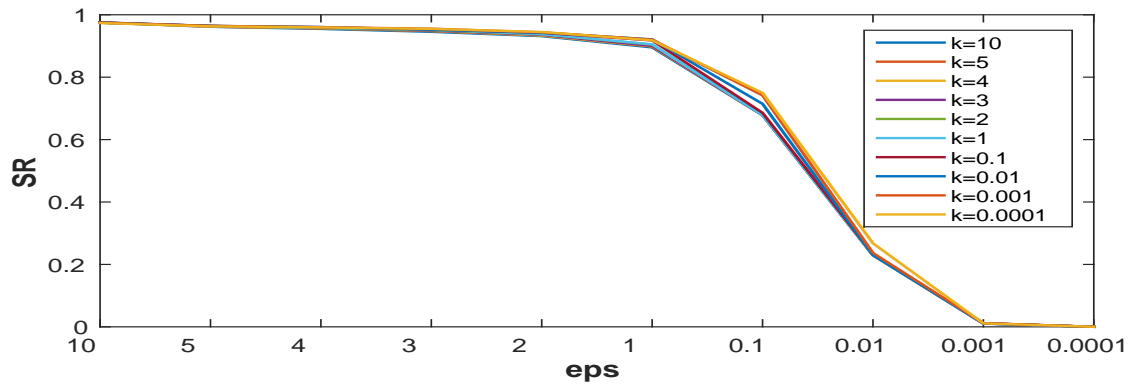
b)  $\mathcal{St}_g$  distributed samples  $x \sim \mathcal{St}_g(x|\alpha, \beta)$



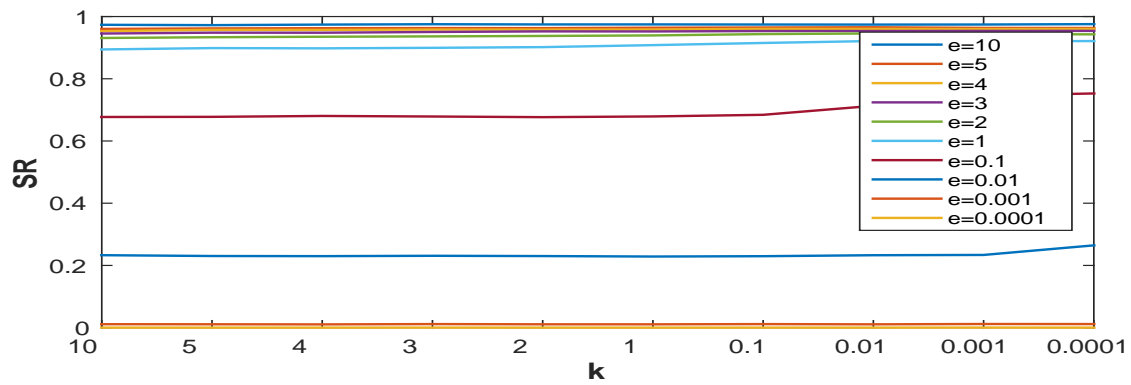
c) Sparse signal obtained by thresholding  $|x| < 0.1$

Figure A.2: Sparse signal, distributed following the generalized Student-t distribution after thresholding

## A.2.2 - Bayesian parameter estimation



a) Sparsity rate as a function of  $\epsilon$ , Eq. (A.30a)



b) Sparsity rate as a function of  $k$ , Eq. (A.30b)

Figure A.3: Behaviour of the Sparsity Rate as a function of  $\epsilon$  for different values of  $k$  Eq. (A.30a)

# B

## Publications

### Journal Articles

---

- [J1] **Mircea Dumitru**, Ali Mohammad-Djafari, Simona Baghai Sain (2016). *Precise periodic components estimation for chronobiological signals through Bayesian Inference with sparsity enforcing prior*. EURASIP Journal on Bioinformatics and Systems Biology, special issue on Bayesian Methods for Computational Systems Biology, 2016, doi:10.1186/s13637-015-0033-6
- [J2] Ali Mohammad-Djafari, **Mircea Dumitru** (2015). *Bayesian sparse solutions to linear inverse problems with non-stationary noise with Student-t priors*, Digital Signal Processing, 2015, doi:10.1016/j.dsp.2015.08.005
- [J3] Xiao-Mei Li, Ali Mohammad-Djafari, **Mircea Dumitru**, Sandrine Dulong, Elisabeth Filipksi, Sandrine Siffroi-Fernandez, Ali Mteyrek, Francesco Scaglione, Catherine Guettier, Franck Delaunay and Francis Lévi (2013). *A circadian clock transcription model for the personalization of cancer chronotherapy*, Cancer Research, American Association for Cancer Research, 2013, 73 (24), pp.7176-88. <10.1158/0008-5472.CAN-13-1528>

### Conference Papers

---

- [C1] Li Wang, Ali Mohammad-Djafari, Nicolas Gac, **Mircea Dumitru**,(2016). Computed tomography reconstruction based on a hierarchical model and variational bayesian method. *In 41st IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), actes : ISBN:— ,in Shanghai, China*

- [C2] **Mircea Dumitru**, Ali Mohammad-Djafari (2015). Periodic components estimation in chronobiological time series via a Bayesian approach. *In 23rd IEEE European Signal Processing Conference, (EUSIPCO 2015), actes : ISBN: 978-0-9928626-4-0, pages —, Nice Congress Center, in Nice, France*
- [C3] **Mircea Dumitru**, Ali Mohammad-Djafari (2014). Estimating the periodic components of a biomedical signal through inverse problem modelling and Bayesian inference with sparsity enforcing prior. *In the 34th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering (MaxEnt 2014) Proceedings, actes : volume 1641, pages 548-555, Château Clos Lucé, Parc Leonardo Da Vinci, Amboise, France*

## Oral and Poster presentations

---

- [OP1] Xiao-Mei Li, **Mircea Dumitru**, Narin Ozturk , Ali Mohammad-Djafari , Francis Lévi (2015). Characterization of circadian rhythms of liver Per2::luc expression in freely moving mice. *European Biological Rhythms Society (EBRS)/World Congress of Chronobiology (WCC) meeting, Manchester, UK*
- [OP2] **Mircea Dumitru**, Ali Mohammad-Djafari, Xiao-Mei Li, Francis Lévi (2013). Prediction of weight loss in cancer treatment from the Reverb- $\alpha$  and Bmal1 gene expressions. *Interdisciplinary Symposium on Signals and Systems for Medical Applications (ISSSMA13), June 2013, Paris France.*
- [OP3] Ali Mohammad-Djafari, Leila Gharsalli, **Mircea Dumitru** (2012). Bayesian variational approximation implementation for linear inverse problem with infinite Gaussian mixture model. *Journées Résolution de problèmes inverses : optimisation et parallélisation (GDR ISIS). Paris France.*

# Bibliography

- [AH06] K. Aas and I.H. Haff. The generalized hyperbolic skew student's t-distribution. *Journal of Financial Econometrics*, 4(2):275–309, 2006.
- [AKZ06] A. Achim, E. E. Kuruoglu, and J. Zerubia. Sar image filtering based on the heavy-tailed rayleigh model. *IEEE Transactions on Image Processing*, 15(9):2686 – 2693, Sept. 2006.
- [AMD09] Hacheme Ayasso and Ali Mohammad-Djafari. Joint image restoration and segmentation using Gauss-Markov-Potts prior models and variational bayesian computation. In *Proceeding of the 15th IEEE International Conference on Image Processing, (ICIP)*, pages 1297–1300, Égypte, 2009.
- [AMD10] Hacheme Ayasso and Ali Mohammad-Djafari. Joint NDT image restoration and segmentation using Gauss–Markov–Potts prior models and variational bayesian computation. *IEEE Transactions on Image Processing*, 19(9):2265–2277, 2010.
- [BCMR09] M. Beaumont, J.M. Cornuet, J.M. Marin, and C.P. Robert. Adaptive approximate bayesian computation. *Biometrika*, 96(4):983–990, 2009.
- [Bea03] M. Beal. *Variational Algorithms for Approximate Bayesian Inference*. PhD thesis, Gatsby Computational Neuroscience Unit, University College London, 2003.
- [BF10] M. Blum and O. François. Non-linear regression models for approximate bayesian computation. *Statistics and Computing*, 20(1):63–73, 2010.
- [Blu10] M. Blum. Approximate bayesian computational: a non-parametric perspective. *Journal of the American Statistical Association*, 491:1178–1187, 2010.
- [CGHM14] N. Chen, D. Giannakis, R. Herbei, and A.J. Majda. An mcmc algorithm for parameter estimation in signals with hidden intermittent instability. *SIAM/ASA Journal on Uncertainty Quantification*, 2:647–669, 2014.
- [CGI96] F Champagnat, Y. Goussard, and J. Idier. Unsupervised deconvolution of sparse spike trains using stochastic approximation. *IEEE Trans. signal processing*, 44(12):2988–2998, 1996.
- [CKS15] Y. Chen, E. E. Kuruoglu, and H. C. So. Optimum linear regression in additive cauchy–gaussian noise. *Signal Processing*, 106:312–318, 2015.

## BIBLIOGRAPHY

---

- [CMD97] H. Carfantan and Ali Mohammad-Djafari. A Bayesian framework for nonlinear diffraction tomography. In *IEEE EURASIP Workshop on Nonlinear Signal and Image Processing NSIP'97*, Mackinac island, MI, USA, September 1997.
- [CW98] T. F. Chan and C. K. Wong. Total variation blind deconvolution. *IEEE Trans Image Process*, 7(3):370–375, 1998.
- [CWB08] E. J. Candès, M. Wakin, and S. Boyd. Enhancing sparsity by reweighted  $\ell_1$  minimization. *Journal of Fourier Analysis and Applications*, 14:877–905, 2008.
- [DB08] Julia A Dobrosotskaya and Andrea L Bertozzi. A wavelet-laplace variational technique for image deconvolution and inpainting. *IEEE Trans Image Process*, 17(5):657–663, May 2008.
- [DDJ09] P. Del Moral, A. Doucet, and A. Jasra. An adaptive sequential monte carlo method for approximate bayesian computation. Technical report, Imperial College, London, 2009.
- [DDM04] I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Appl. Math*, 57:1413–1457, 2004.
- [Dee81] T.J. Deeming. Deconvolution and reflection coefficient estimation using a generalized minimum entropy principle. *Proc. 51st Ann. Meeting of the Soc. of Exploration Geophysicists*, page 1, 1981.
- [Don81] D.L. Donoho. On minimum entropy deconvolution. *Applied Time Series Analysis II*, page 1, 1981.
- [DR85] G. Demoment and R. Reynaud. Fast minimum-variance deconvolution. *IEEE Transactions on Acoustics Speech and Signal Processing*, ASSP-33:1324–1326, 1985.
- [DSA10] C. Dibner, U. Schibler, and U. Albrecht. The mammalian circadian timing system: organisation and coordination of central and peripheral clocks. *Annu. Rev. Physiol.*, 72:517–549, 2010.
- [EIM90] Kane. S. E., Pastan. I., and Gottesman. M. M. Genetic basis of multidrug resistance of tumor cells. *J Bionenerg Biomembr*, 22:593–618, 1990.
- [F.01] Levi F. Circadian chronotherapy for human cancers. *Lancet Oncol*, 2:307–315, 2001.
- [FDMD05] Olivier Féron, Bernard Duchêne, and Ali Mohammad-Djafari. Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data. *Inverse Problems*, 21(6):95–115, Dec 2005.
- [FDMD07] Olivier Féron, Bernard Duchêne, and Ali Mohammad-Djafari. Microwave imaging of piecewise constant objects in a 2D-TE configuration. *International Journal of Applied Electromagnetics and Mechanics*, 26(6):167–174, IOS Press 2007.

- [Fer99] C Fernandez. Multivariate Student-t regression models: Pitfalls and inference. *Biometrika*, 86(1):153–167, 1999.
- [Fio04] Simone Fiori. Fast fixed-point neural blind-deconvolution algorithm. *IEEE Trans Neural Netw*, 15(2):455–459, Mar 2004.
- [Fis25] R.A. Fisher. Applications of students, distribution. *Metron*, 5:90–104, 1925.
- [FIW<sup>+</sup>05] E. Filipski, P. F. Innominato, M. Wu, X. M. Li, and S. Iacobelli. Effects of light and food schedules on liver and tumor molecular clocks in mice. *J. Natl. Cancer Inst.*, 97:507–517, 2005.
- [FKL<sup>+</sup>02] E. Filipski, V. M. King, X. Li, T. G. Granda, and M. C. Mormont. Host circadian clock as a control point in tumor progression. *J. Natl. Cancer Inst.*, 94:690–697, 2002.
- [FL03] L. Fu and C. C. Lee. The circadian clock: pacemaker and tumor suppressor. *Nat. Rev. Cancer*, 3:350–361, 2003.
- [GCSR03] Andrew Gelman, John B. Carlin, Hal S. Stern, and Donald B. Rubin. *Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science)*. Chapman and Hall/CRC, 2 edition, July 2003.
- [GDI90] Y. Goussard, G. Demoment, and J. Idier. A new algorithm for iterative deconvolution of sparse spike trains. *Proc. Int. Conf. ASSP*, page 3, 1990.
- [GG84] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell*, 6(6):721–741, Jun 1984.
- [GHBMD13] L. Gharsalli, A. Hacheme, D. Bernard, and A. Mohammad-Djafari. Microwave tomography for breast cancer detection within a variational bayesian approach. In *IEEE European Signal Processing Conference, EUSIPCO*, volume nc, page nc, Marrakech, Morocco, 2013.
- [GILC08] D. Ge, J. Idier, and E. Le Carpentier. A new mcmc algorithm for blind bernoulli- gaussian deconvolution. In *Proceedings of EUSIPCO: Septembre 2008; Lausanne, Suisse*, 2008.
- [GLRI93] G. Gindi, M. Lee, A. Rangarajan, and Zubal I. Bayesian reconstruction of functional images using anatomical information as priors. *ieeemi*, MI-12(4):670–680, 1993.
- [Gos08] W. S. Gosset. The probable error of a mean. *Biometrika*, 6, 1908.
- [GP93] M. M. Gottesman and I. Pastan. Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annu Rev Biochem*, 62:385–427, 1993.
- [GRS96] W.R. Gilks, S. Richardson, and D.J. Spiegelhalter. *Monte Carlo Markov Chains in practice*. Chapman and Hall, 1996.
- [Had01] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton Univ. Bull.*, 13, 1901.



## BIBLIOGRAPHY

---

- [Haz01] M. Hazewinkel. *Student distribution, Encyclopedia of Mathematics*. Springer, 2001.
- [Hel75] F. R. Helmert. Über die bestimmung des wahrscheinlichen fehlers aus einer endlichen anzahl wahrer beobachtungsfehler. *Z. Math.*, 20:300–303, 1875.
- [Hel76a] F. R. Helmert. Die genauigkeit der formel von peters zur berechnung des wahrscheinlichen beobachtungsfehlers directer beobachtungen gleicher genauigkeit. *Astron. Nachr.*, 88:113–132, 1876.
- [Hel76b] F. R. Helmert. Über die wahrscheinlichkeit der potenzsummen der beobachtungsfehler und uber einige damit in zusammenhang stehende fragen. *Z. Math. Phys.*, 21:192–218, 1876.
- [HMC02] R. V. Hogg, J. McKean, and A. T. Craig. *Introduction to Mathematical Statistics (7th Edition)*. Pearson, 2002.
- [Hol92] T. J. Holmes. Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach. *J Opt Soc Am A*, 9(7):1052–1061, Jul 1992.
- [HRM08] M. E. Hastings, A. B. Reddy, and E. S. Maywood. A clockwork web: circadian timing in brain and periphery, in health and disease. *Nat. Rev. Neurosci.*, 4(8):649–661, 2008. cited By 555.
- [HW83] Kenneth M. Hanson and G. W. Wechsung. Bayesian approach to limited-angle reconstruction in computed tomography. 73:1501–1509, November 1983.
- [IFG<sup>+</sup>09] P. F. Innominato, C. Focan, T. Gorlia, T. Moreau, and 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 Research*, 69:4700–4707, 2009.
- [IG90] J. Idier and Y. Goussard. Markov modeling for bayesian multi-channel deconvolution. *Proceedings of IEEE ICASSP*, page 2, 1990.
- [IGB<sup>+</sup>12] P. F. Innominato, S. Giacchetti, G. A. Bjarnason, C. Focan, C. Garufi, and B. Coudert. Prediction of overall survival through circadian rest-activity monitoring during chemotherapy for metastatic colorectal cancer. *Int J Cancer*, 131:2684–2692, 2012.
- [IGM<sup>+</sup>11] P. F. Innominato, S. Giacchetti, T. Moreau, R. Smaaland, C. Focan, and G. A. 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.
- [JSM<sup>+</sup>02] Stuart Jefferies, Kathy Schulze, Charles Matson, Kurt Stoltenberg, and E. Keith Hege. Blind deconvolution in optical diffusion tomography. *Opt Express*, 10(1):46–53, Jan 2002.
- [KBJ00] S. Kotz, N. Balakrishnan, and N.L. Johnson. *Continuous multivariate distributions, models and applications*, volume 334. Wiley-Interscience, 2000.

- [KMS<sup>+</sup>12a] F. Krzakala, M. Mézard, F. Sausset, Y. Sun, and L. Zdeborová. Probabilistic reconstruction in compressed sensing: algorithms, phase diagrams, and threshold achieving matrices. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(08):P08009, 2012.
- [KMS<sup>+</sup>12b] F. Krzakala, M. Mézard, F. Sausset, Y. F. Sun, and L. Zdeborová. Statistical-physics-based reconstruction in compressed sensing. *Phys. Rev. X*, 2:021005, May 2012.
- [KTB04] EE Kuruoglu, A. Tonazzini, and L. Bianchi. Source separation in noisy astrophysical images modelled by markov random fields. In *Image Processing, 2004. ICIP'04. 2004 International Conference on*, volume 4, pages 2701–2704. IEEE, 2004.
- [L76] J. Lüroth. Vergleichung von zwei werten des wahrscheinlichen fehlers. *Astron. Nachr.*, 87:209–220, 1876.
- [LB13] Thibault Lelore and Frédéric Bouchara. Fair: a fast algorithm for document image restoration. *IEEE Trans Pattern Anal Mach Intell*, 35(8):2039–2048, Aug 2013.
- [Li57] J.C.R. Li. Student’s t-distribution. In *Introduction to statistical inference.*, pages 87–104. Edwards Brothers, 1957.
- [LMB<sup>+</sup>09] Gu. L., Tsark. W. M., D. A. Brown, S. Blanchard, T. W. Synold, and S. E. Kane. A new model for studying tissue-specific mdrla gene expression in vivo by live imaging. *Proc Natl Acad Sci USA*, 106:5394–5399, 2009.
- [LMDD<sup>+</sup>13] X.-M. Li, A. Mohammad-Djafari, M. Dumitru, S. Dulong, E. Filipski, S. Siffroi-Fernandez, A. Mteyrek, F. Scaglione, C. Guettier, F. Delaunay, and F. Levi. A circadian clock transcription model for the personalization of cancer chronotherapy. *Cancer Research*, 73(24):7176–7188, 2013. cited By 5.
- [LN11] Haiyong Liao and Michael K Ng. Blind deconvolution using generalized cross-validation approach to regularization parameter estimation. *IEEE Trans Image Process*, 20(3):670–680, Mar 2011.
- [LOD<sup>+</sup>10] F. Levi, A. Okyar, S. Dulong, P. F. Innominato, and J. Clairambault. Circadian timing in cancer treatment. *Annu Rev Pharmacol Toxicol*, 50:377–421, 2010.
- [LS07] F. Levi and U. Schibler. Circadian rhythms: mechanisms and therapeutic implications. *Annu. Rev. Pharmacol. Toxicol.*, 47:593–628, 2007.
- [MD12] A Mohammad-Djafari. Bayesian approach with prior models which enforce sparsity in signal and image processing. *EURASIP Journal on Advances in Signal Processing*, Special issue on Sparse Signal Processing:2012:52, 2012.
- [MDD88] Ali Mohammad-Djafari and G. Demoment. Image restoration and reconstruction using entropy as a regularization functional. *Maximum Entropy and Bayesian Methods in Science and Engineering*, 2:341–355, 1988.
- [MGT12] J. A. Mohawk, C. B. Green, and J. S. Takahashi. Central and peripheral circadian clocks in mammals. *Annu. Rev. Neurosci.*, 35:445–462, 2012.

## BIBLIOGRAPHY

---

- [ML03] M. C. Mormont and F. Levi. Cancer chronotherapy: principles, applications, and perspectives. *Cancer*, 97:155–169, 2003.
- [Mor08] R. G. Mortimer. *Mathematics for Physical Chemistry*. Academic Press. 3 edition., 1908.
- [MPRR11] J.M. Marin, P. Pudlo, C.P. Robert, and R. Ryder. Approximate bayesian computational methods. *arxiv:1101.0955*, 2011.
- [MS04] R. E. Mistlberger and D. J. Skene. Social influences on mammalian circadian rhythms: animal and human studies. *Biol. Rev.*, 79:533–556, 2004.
- [MSC99] N.D.A. Mascarenhas, C.A.N. Santos, and P.E. Cruvinel. Transmission tomography under poisson noise using the anscombe transformation and wiener filtering of the projections. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 423(2):265–271, 1999.
- [MSSM11] Andrés G Marrugo, Michal Sorel, Filip Sroubek, and María S Millán. Retinal image restoration by means of blind deconvolution. *J Biomed Opt*, 16(11):116016, Nov 2011.
- [MWB<sup>+</sup>00] M. C. Mormont, J. Waterhouse, P. Bleuzen, S. Giacchetti, and 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.
- [NMD94] M.K. Nguyen and A. Mohammad-Djafari. Bayesian approach with the maximum entropy priors in image reconstruction from the microwave scattered field data. *ieeMI*, 1994.
- [NMD96] M. Nikolova and Ali Mohammad-Djafari. Eddy current tomography using a binary Markov model. *Signal Processing*, 49:119–132, 1996.
- [NPQ00] M. K. Ng, R. J. Plemmons, and S. Qiao. Regularization of rif blind image deconvolution. *IEEE Trans Image Process*, 9(6):1130–1134, 2000.
- [OPA<sup>+</sup>11] A. Okyar, E. Piccolo, C. Ahowesso, E. Filipski, V. Hossard, C. Guettier, R. La Sorda, N. Tinari, S. Iacobelli, and F. Levi. Strain- and sex-dependent circadian changes in *abcc2* transporter expression: implications for irinotecan chronotolerance in mouse ileum. *PLoS One*, 6, 2011.
- [OTMB<sup>+</sup>13] E. Ortiz-Tudela, A. Mteyrek, A. Ballesta, P. F. Innominato, and F. Levi. Cancer chronotherapeutics: experimental, theoretical, and clinical aspects. *Handb Exp Pharmacol.*, 217:261–288, 2013.
- [PB13] Hanjie Pan and Thierry Blu. An iterative linear expansion of thresholds for  $l_1$ -based image restoration. *IEEE Trans Image Process*, 22(9):3715–3728, Sep 2013.
- [PCS11] J.T. Parker, V. Cevher, and P. Schniter. Compressive sensing under matrix uncertainties: An approximate message passing approach. In *Signals, Systems and Computers (ASILOMAR), 2011 Conference Record of the Forty Fifth Asilomar Conference on*, pages 804–808, nov. 2011.

- [PL99] S. U. Pillai and B. Liang. Blind image deconvolution using a robust gcd approach. *IEEE Trans Image Process*, 8(2):295–301, 1999.
- [PS10] N. Polson and J. Scott. Shrink globally, act locally: sparse Bayesian regularization and prediction. *Bayesian Statistics 9*, 2010.
- [RMC09] H. Rue, S. Martino, and N. Chopin. Approximate bayesian inference for latent gaussian models using integrated nested laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B*, 71(2):319–392, 2009.
- [RMW12] Mohammad Rostami, Oleg Michailovich, and Zhou Wang. Image deblurring using derivative compressed sensing for optical imaging application. *IEEE Trans Image Process*, 21(7):3139–3149, Jul 2012.
- [SLC<sup>+</sup>13] C. Saini, A. Liani, T. Curie, P. Gos, F. Kreppel, Y. Emmenegger, L. Bonacina, J. P. Wolf, Y. A. Poget, P. Franken, and 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.
- [SR94] S. Senn and W. Richardson. The first t-test. *Statistics in Medicine*, 13:785–803, 1994.
- [SR12] Philip Schniter and Sundeep Rangan. Compressive Phase Retrieval via Generalized Approximate Message Passing. In *Proceedings of Allerton Conference on Communication, Control, and Computing*, 2012.
- [SSKS00] S. E. Sephton, R. M. Sapolsky, H. C. Kraemer, and D. Spiegel. Diurnal cortisol rhythm as a predictor of breast cancer survival. *J. Natl. Cancer Inst.*, 92:994–1000, 2000.
- [SSL<sup>+</sup>11] C. Saini, D. M. Suter, A. Liani, P. Gos, and U. Schibler. The mammalian circadian timing system: synchronization of peripheral clocks. *Cold Spring Harbor Symposia on Quantitative Biology*, 76:39–47, 2011.
- [TGS06] J. A. Tropp, A. C. Gilbert, and M. J. Strauss. Algorithms for simultaneous sparse approximation. part i: Greedy pursuit. *Signal Processing, special issue "sparse approximations in signal and image processing"*, 86:572–588, April 2006.
- [TH94] Y. Touitou and E. Haus. Biologic rhythms in clinical and laboratory medicine. *Springer-Verlag*, 1994.
- [Tip01] M. Tipping. Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211–244, 2001.
- [Tro06] J. A. Tropp. Algorithms for simultaneous sparse approximation. part ii: Convex relaxation. *Signal Processing, special issue "Sparse approximations in signal and image processing"*, 86:589–602, April 2006.
- [VR02] W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S*. Springer, 2002.

## BIBLIOGRAPHY

---

- [Wah82] G. Wahba. Constrained regularization for ill-posed linear operator equations with applications in meteorology and medicine. *in Statistical Decision Theory and Related Topics III New-York: Academic*, pages 383–417, 1982.
- [WBJ05] John Winn, Christopher M. Bishop, and Tommi Jaakkola. Variational message passing. *Journal of Machine Learning Research*, 6:661–694, 2005.
- [WHmPw08] Liang Wei, Lei Hua-ming, and Que Pei-wen. Sparsity enhancement for blind deconvolution of ultrasonic signals in nondestructive testing application. *Rev Sci Instrum*, 79(1):014901, Jan 2008.
- [Wig77] R.A. Wiggins. Minimum entropy deconvolution. *IEEE Proceedings Intern. Symp. Computer-Aided Seismic Analysis and Discrimination*, 1977.
- [WMMY02] R. Walpole, R. Myers, S. Myers, and K. Ye. *Probability and Statistics for Engineers and Scientists*. Pearson Education, 7th edition, 2002.
- [YBW85] R. Yarlagadda, J.B. Bednar, and T.L. Watt. Fast algorithms for lp deconvolution. *IEEE Transactions on Acoustics Speech and Signal Processing*, ASSP-33, 1985.
- [YFZ12] Luxin Yan, Houzhang Fang, and Sheng Zhong. Blind image deconvolution with spatially adaptive total variation regularization. *Opt Lett*, 37(14):2778–2780, Jul 2012.
- [YLZF12] Luxin Yan, Hai Liu, Sheng Zhong, and Houzhang Fang. Semi-blind spectral deconvolution with adaptive tikhonov regularization. *Appl Spectrosc*, 66(11):1334–1346, Nov 2012.
- [Zel76] Arnold Zellner. Bayesian and Non-Bayesian Analysis of the Regression Model with Multivariate Student-t Error Terms. *J. of American Stat. Assoc.*, 71(354):400–405, 1976.
- [ZL11] Wangmeng Zuo and Zhouchen Lin. A generalized accelerated proximal gradient approach for total-variation-based image restoration. *IEEE Trans Image Process*, 20(10):2748–2759, Oct 2011.

# List of Figures

1.1	Exemple de signaux chronobiologiques obtenus dans des expériences de traitement du cancer: le	
1.2	Les données brutes pour un signal representent des modelés de repos-activité en chronobiologie:	
1.3	Les données brutes pour un signal representent l'expression des gènes en chronobiologie: avant le	
1.4	Densité de la loi Normale et densité de la loi de Student. . . . .	15
1.5	La distribution bivarié Normale - Inverse Gamma modélisant $(f_{kj}, v_{f_j})$ et la distribution margina	
1.6	Les equations de mise à jour des paramètres pour l'estimation MAP. . . . .	18
1.7	Les equations de mise à jour des paramètres pour l'estimation EAP via ABV, séparabilité partielle	
1.8	Les equations de mise à jour des paramètres pour l'estimation EAP via ABV, séparabilité totale.	
2.1	Example of chronobiological signals obtained in cancer treatment experiments: the rest-activity s	
2.2	Raw data for a rest-activity pattern signal in chronobiology: before treatment (left, 3 days length)	
2.3	Raw data for a gene expression signal in chronobiology: before treatment (left, 3 days length), du	
3.1	Synthetic data: sparse theoretical PC amplitudes vector $\mathbf{f}$ with 3 non-zero peaks corresponding to	
3.2	Estimated PC amplitudes vectors $\widehat{\mathbf{f}}_0$ (3.2c) and $\widehat{\mathbf{f}}$ (3.2d) for the original signal $\mathbf{g}$ (3.2a) and the no	
3.3	Estimated PC amplitudes vectors $\widehat{\mathbf{f}}_0$ (3.3a) and $\widehat{\mathbf{f}}$ (3.3b). Periods are not scaled in the plot. 31	
3.4	Real data: A 4 days length signal recorded in cancer treatment experiments (3.4a), the PC amplitu	
3.5	Estimated PC amplitudes vectors $\widehat{\mathbf{f}}_0$ (3.5c) and $\widehat{\mathbf{f}}$ (3.5d) for the zero padded original signal $\mathbf{g}$ (3.5	
3.6	Estimated PC amplitudes vectors $\widehat{\mathbf{f}}_0$ (3.6a) and $\widehat{\mathbf{f}}$ (3.6b), corresponding to zero padded signals. P	
3.7	Estimated PC amplitudes vectors $\widehat{\mathbf{f}}_0$ (3.7c) and $\widehat{\mathbf{f}}$ (3.7d) for the zero padded original signal $\mathbf{g}$ (3.7	
3.8	Zero padded signal (a) and the corresponding FFT PC amplitudes vector. . . .	34
3.9	Direct Inversion: Signal without noise . . . . .	35
3.10	Direct Inversion: Signal with noise, 50dB . . . . .	36
3.11	Direct Inversion: Signal with noise, 20dB . . . . .	36
3.12	Direct Inversion: Signal with noise, 05dB . . . . .	37
3.13	Direct Inversion $L_1$ and $L_2$ errors: PC estimation errors and signal reconstruction errors	38

## LIST OF FIGURES

---

3.14	Least Squares: Signal without noise . . . . .	38
3.15	Least Squares: Signal with noise, 15dB . . . . .	39
3.16	Least Squares: Signal with noise, 10dB . . . . .	40
3.17	Least Squares: Signal with noise, 05dB . . . . .	41
3.18	Least Squares $L_1$ and $L_2$ errors: PC estimation errors and signal reconstruction errors	41
3.19	Regularized Least Squares, $\lambda = 0.1$ : Signal without noise . . . . .	42
3.20	Regularized Least Squares, $\lambda = 0.1$ : Signal with noise, 5dB . . . . .	42
3.21	Regularized Least Square $L_1$ and $L_2$ errors: PC estimation and signal reconstruction	43
3.22	Regularized Least Square $\delta \mathbf{g}_1$ and $\delta \mathbf{g}_2$ signal reconstruction errors for different values of $\lambda$	44
3.23	Regularized Least Square $\delta \mathbf{f}_1$ and $\delta \mathbf{f}_2$ PC estimation errors for different values of $\lambda$	44
3.24	LASSO: Signal without noise . . . . .	45
3.25	LASSO: Signal with noise, 10dB . . . . .	46
3.26	LASSO: Signal with noise, 5dB - Accurate estimation . . . . .	46
3.27	LASSO: Signal with noise, 5dB - Inaccurate estimation . . . . .	47
3.28	LASSO $L_1$ and $L_2$ errors stability: PC estimation and signal reconstruction - 10dB	48
3.29	LASSO $L_1$ and $L_2$ errors stability: PC estimation and signal reconstruction - 5dB	48
3.30	LASSO $L_1$ and $L_2$ errors stability: PC estimation and signal reconstruction - 5dB	49
3.31	LASSO $L_1$ and $L_2$ errors: PC estimation and signal reconstruction . . . . .	49
4.1	Normal and Student-t distribution . . . . .	60
4.2	Normal vs. Student-t distribution . . . . .	60
4.3	Normal - Inverse Gamma distribution $(f_{kj}, v_{fj})$ and the corresponding marginal $f_{kj}   \alpha_{f0}, \beta_{f0}$	61
4.4	Normal vs. Normal - Inverse Gamma marginal distributions . . . . .	63
4.5	Updating scheme: JMAP estimation for generalized Student-t prior model . . . . .	65
4.6	Updating scheme: PM estimation via VBA (partial separability) for generalized Student-t prior model	
4.7	Inverse Gamma distributions for different $\alpha, \beta$ parameters . . . . .	72
4.8	Sparsity mechanism: the sparsity is imposed (also) via the corresponding variance	72
4.9	$x \sim \mathcal{IG}(x \alpha, \beta)$ ; $E[x] = 0.1$ ; $\text{Var}[x] = 1, 0.1, 0.01$ and $0.001$ . . . . .	73
4.10	$x \sim \mathcal{IG}(x \alpha, \beta)$ ; $E[x] = 0.01$ ; $\text{Var}[x] = 1, 0.1, 0.01$ and $0.001$ . . . . .	73
4.11	$x \sim \mathcal{IG}(x \alpha, \beta)$ ; $E[x] = 0.001$ ; $\text{Var}[x] = 1, 0.1, 0.01$ and $0.001$ . . . . .	74
4.12	Different Generalized Student-t distributions for which the means $E_{St_g} = 0$ and the variances $\text{Var}_{St_g} =$	
4.13	Generalized Student-t: $\text{Var}_{St_g} = 0.1$ ; $\text{Var}_{\mathcal{IG}} = 0.1, 0.01, 0.001$ ; . . . . .	75
4.14	Different Generalized Student-t distributions for which the means $E_{St_g} = 0$ and the variances $\text{Var}_{St_g} =$	

---

4.15 Generalized Student-t: $\text{Var}_{St_g} = 0.01$ ; $\text{Var}_{IG} = 0.1, 0.01, 0.001$ ; . . . . .	75
4.16 Different Generalized Student-t distributions for which the means $E_{St_g} = 0$ and the variances $\text{Var}_{St_g}$	
4.17 Generalized Student-t: $\text{Var}_{St_g} = 0.001$ ; $\text{Var}_{IG} = 0.1, 0.01, 0.001$ ; . . . . .	76
4.18 Updating scheme: PM estimation via VBA (full separability) for generalized Student-t prior mod	
5.1 $\mathbf{f}$ PC vector, theoretical signal $\mathbf{g}_0$ and input signal $\mathbf{g} = \mathbf{g}_0 + \boldsymbol{\epsilon}$ of the model (5dB)	81
5.2 JMAP IGSM Estimation (5dB) . . . . .	82
5.3 JMAP IGSM $L_2$ error measured for 10 different noise realisations (5dB) . . .	83
5.4 PM (via VBA, partial separability) IGSM Estimation (5dB) . . . . .	84
5.5 PM (via VBA, partial separability) IGSM hyperparameters and $\mathbf{f}$ convergency	85
5.6 PM (via VBA, partial separability) IGSM $L_2$ error measured for 10 different noise realisations (5	
5.7 PM (via VBA, full separability) IGSM Estimation (5dB) . . . . .	87
5.8 PM (via VBA, full separability) IGSM $L_2$ error measured for 10 different noise realisations (5dB	
5.9 Methods Comparison (5dB) . . . . .	89
5.10 $L_2$ Errors estimation (5dB) . . . . .	90
5.11 $\mathbf{f}$ PC vector, theoretical signal $\mathbf{g}_0$ and input signal $\mathbf{g} = \mathbf{g}_0 + \boldsymbol{\epsilon}$ of the model (10dB)	91
5.12 JMAP IGSM Estimation (10dB) . . . . .	92
5.13 JMAP IGSM $L_2$ error measured for 10 different noise realisations (10dB) . .	93
5.14 PM (via VBA, partial separability) IGSM Estimation (10dB) . . . . .	94
5.15 PM (via VBA, partial separability) IGSM hyperparameters and $\mathbf{f}$ convergency	94
5.16 PM (via VBA, partial separability) IGSM $L_2$ error measured for 10 different noise realisations (1	
5.17 PM (via VBA, full separability) IGSM Estimation (10dB) . . . . .	96
5.18 PM (via VBA, full separability) IGSM $L_2$ error measured for 10 different noise realisations (10dB	
5.19 Methods Comparison (10dB) . . . . .	98
5.20 $L_2$ Errors estimation (10dB) . . . . .	99
5.21 $\mathbf{f}$ PC vector, theoretical signal $\mathbf{g}_0$ and input signal $\mathbf{g} = \mathbf{g}_0 + \boldsymbol{\epsilon}$ of the model (15dB)	100
5.22 JMAP IGSM Estimation (15dB) . . . . .	101
5.23 JMAP IGSM $L_2$ error measured for 10 different noise realisations (15dB) . .	102
5.24 PM (via VBA, partial separability) IGSM Estimation (15dB) . . . . .	103
5.25 PM (via VBA, partial separability) IGSM hyperparameters and $\mathbf{f}$ convergency	103
5.26 PM (via VBA, partial separability) IGSM $L_2$ error measured for 10 different noise realisations (1	
5.27 PM (via VBA, full separability) IGSM Estimation (15dB) . . . . .	105
5.28 PM (via VBA, full separability) IGSM $L_2$ error measured for 10 different noise realisations (15dB	

---



## LIST OF FIGURES

---

5.29	Methods Comparison (15dB) . . . . .	107
5.30	$L_2$ Errors estimation (15dB) . . . . .	108
5.31	Synthetic simulation data (05dB): sine PC amplitudes vector $f_1$ , (5.31a), cosine PC amplitudes vector	
5.32	Synthetic simulation (05dB): comparison between the Proposed Method estimations and the synthetic	
5.33	Synthetic simulation (05dB): comparison between the LASSO estimations and the synthetic data	111
5.34	Stability: L1 and L2 reconstruction errors (05dB): $\delta g_1$ and $\delta g_2$ : Proposed Method vs. LASSO	112
5.35	Stability: L1 and L2 estimation errors (05dB): $\delta f_1$ and $\delta f_2$ : Proposed Method vs. LASSO	113
5.36	Synthetic simulation data (10dB): sine PC amplitudes vector $f_1$ , (5.36a), cosine PC amplitudes vector	
5.37	Synthetic simulation (10dB): comparison between the Proposed Method estimations and the synthetic	
5.38	Synthetic simulation (10dB): comparison between the LASSO estimations and the synthetic data	116
5.39	Stability: L1 and L2 reconstruction errors (10dB): $\delta g_1$ and $\delta g_2$ : Proposed Method vs. LASSO	117
5.40	Stability: L1 and L2 estimation errors (10dB): $\delta f_1$ and $\delta f_2$ : Proposed Method vs. LASSO	118
5.41	L1 and L2 estimation errors: $\delta g_1$ and $\delta g_2$ : Proposed Method vs. LASSO . . .	119
5.42	L1 and L2 estimation errors: $\delta f_1$ and $\delta f_2$ : Proposed Method vs. LASSO . .	120
6.1	<b>CT 502, A1 Mouse:</b> Activity (a) and Photon Absorption (b) raw data . . . .	122
6.2	Activity raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	123
6.3	Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	123
6.4	PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	124
6.5	PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	124
6.6	PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	125
6.7	DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	125
6.8	DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	125
6.9	PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during	126
6.10	PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	126
6.11	PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	126
6.12	DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	127
6.13	Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	128
6.14	Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	128
6.15	PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	128
6.16	PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	129
6.17	PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	129
6.18	DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	129

---

6.19 DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	130
6.20 PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during	130
6.21 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	130
6.22 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	130
6.23 DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	130
6.24 <b>CT 502, A2 Mouse</b> : Activity (a) and Photon Absorption (b) raw data . . . . .	132
6.25 Activity raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	133
6.26 Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	133
6.27 PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	134
6.28 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	134
6.29 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	134
6.30 DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	134
6.31 DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	135
6.32 PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during	135
6.33 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	135
6.34 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	136
6.35 DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	136
6.36 Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	137
6.37 Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	137
6.38 PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	138
6.39 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	138
6.40 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	138
6.41 DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	139
6.42 DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	139
6.43 PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during	139
6.44 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	140
6.45 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	140
6.46 DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	140
6.47 <b>CT 502, A3 Mouse</b> : Activity (a) and Photon Absorption (b) raw data . . . . .	142
6.48 Activity raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	143
6.49 Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	143
6.50 PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	144

---

## LIST OF FIGURES

---

6.51	PC Stability: Proposed Method (a) vs. FFT (b)	144
6.52	PC Stability: Proposed Method (a) vs. FFT (b), normalized results	144
6.53	DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	144
6.54	DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	145
6.55	PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during	145
6.56	PC Stability: Proposed Method (a) vs. FFT (b)	145
6.57	PC Stability: Proposed Method (a) vs. FFT (b), normalized results	146
6.58	DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	146
6.59	Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	146
6.60	Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)	147
6.61	PC Stability: PC estimation via FFT and VBA for 4-days length signals	148
6.62	PC Stability: Proposed Method (a) vs. FFT (b)	148
6.63	PC Stability: Proposed Method (a) vs. FFT (b), normalized results	148
6.64	DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	148
6.65	DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	149
6.66	PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during	149
6.67	PC Stability: Proposed Method (a) vs. FFT (b)	149
6.68	PC Stability: Proposed Method (a) vs. FFT (b), normalized results	150
6.69	DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	150
6.70	<b>CT 502, B1 Mouse:</b> Activity (a) and Photon Absorption (b) raw data	151
6.71	Activity raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	152
6.72	Considered signal (a) and the corresponding PC via VBA (b) and FFT (c)	152
6.73	PC Stability: PC estimation via FFT and VBA for 4-days length signals	153
6.74	PC Stability: Proposed Method (a) vs. FFT (b)	153
6.75	PC Stability: Proposed Method (a) vs. FFT (b), normalized results	154
6.76	DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	154
6.77	DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	154
6.78	PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during	155
6.79	PC Stability: Proposed Method (a) vs. FFT (b)	155
6.80	PC Stability: Proposed Method (a) vs. FFT (b), normalized results	155
6.81	DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c)	156
6.82	Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	157

---

6.83 Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	157
6.84 PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	157
6.85 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	158
6.86 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	158
6.87 DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	158
6.88 DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	159
6.89 PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during	159
6.90 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	159
6.91 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	160
6.92 DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	160
6.93 <b>CT 502, B2 Mouse</b> : Activity (a) and Photon Absorption (b) raw data . . . . .	162
6.94 Activity raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	162
6.95 Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	163
6.96 PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	163
6.97 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	164
6.98 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	164
6.99 DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	164
6.100 Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	165
6.101 Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	166
6.102 PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	166
6.103 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	167
6.104 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	167
6.105 DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	167
6.106 <b>CT 502, B3 Mouse</b> : Activity (a) and Photon Absorption (b) raw data . . . . .	168
6.107 Activity raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	168
6.108 Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	169
6.109 PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	170
6.110 PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	170
6.111 PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	170
6.112 DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	171
6.113 DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	171
6.114 PC Stability: PC estimation via FFT and VBA for 4-days length signals, Activity DD, during	171

---

## LIST OF FIGURES

---

6.115	PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	172
6.116	PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	172
6.117	DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	172
6.118	Photon raw data (a) and the corresponding parts,(b),(c),(d),normalized and one-hour sampled	173
6.119	Considered signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	174
6.120	PC Stability: PC estimation via FFT and VBA for 4-days length signals . . .	174
6.121	PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	175
6.122	PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	175
6.123	DD <b>before</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	175
6.124	DD <b>during</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . .	176
6.125	PC Stability: PC estimation via FFT and VBA for 4-days length signals, Photon DD, during	176
6.126	PC Stability: Proposed Method (a) vs. FFT (b) . . . . .	176
6.127	PC Stability: Proposed Method (a) vs. FFT (b), normalized results . . . . .	177
6.128	DD <b>after</b> signal (a) and the corresponding PC via VBA (b) and FFT (c) . . .	177
A.1	Normal-Inverse Gamma distribution (a) and the corresponding marginal $St_g$ (b). Illustrated for $\epsilon = 0$ .	
A.2	Sparse signal, distributed following the generalized Student-t distribution after thresholding	193
A.3	Behaviour of the Sparsity Rate as a function of $\epsilon$ for different values of k Eq. (A.30a)	194

# List of Tables

4.1	Mean and variance for $\mathcal{S}t_g$ via $\epsilon$ and $w$ . . . . .	74
6.1	CT 502 A1: Activity Dominant Period Stability: Proposed Method vs. FFT . . . . .	127
6.2	CT 502 A1: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT	131
6.3	CT 502 A2: Activity Dominant Period Stability: Proposed Method vs. FFT . . . . .	136
6.4	CT 502 A2: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT	141
6.5	CT 502 A3: Activity Dominant Period Stability: Proposed Method vs. FFT . . . . .	147
6.6	CT 502 A3: Photon Dominant Period Stability: Proposed Method vs. FFT . . . . .	151
6.7	CT 502 B1: Activity Dominant Period Stability: Proposed Method vs. FFT . . . . .	156
6.8	CT 502 B1: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT	160
6.9	CT 502 B2: Activity Dominant Period Stability: Proposed Method vs. FFT . . . . .	165
6.10	CT 502 B2: Photon Absorption Dominant Period Stability: Proposed Method vs. FFT	168
6.11	CT 502 B3: Activity Dominant Period Stability: Proposed Method vs. FFT . . . . .	173
6.12	CT 502 B3: Photon Dominant Period Stability: Proposed Method vs. FFT . . . . .	178

**Titre :** Approche bayésienne de l'estimation des composantes périodiques des signaux en chronobiologie

**Mots clés :** Estimation de composantes périodiques, Problèmes inverses, Approches bayésiennes, Modèle hiérarchique, Renforcement de parcimonie, Chronobiologie

**Résumé :** La toxicité et l'efficacité de plus de 30 agents anticancéreux présentent de très fortes variations en fonction du temps de dosage. Par conséquent, les biologistes qui étudient le rythme circadien ont besoin d'une méthode très précise pour estimer le vecteur de composantes périodiques (CP) de signaux chronobiologiques. En outre, dans les développements récents, non seulement la période dominante ou le vecteur de CP présentent un intérêt crucial, mais aussi leurs stabilités ou variabilités. Dans les expériences effectuées en traitement du cancer, les signaux enregistrés correspondant à différentes phases de traitement sont courts, de sept jours pour le segment de synchronisation jusqu'à deux ou trois jours pour le segment après traitement. Lorsque'on étudie la stabilité de la période dominante nous devons considérer des signaux très court par rapport à la connaissance a priori de la période dominante, placée dans le

domaine circadien. Les approches classiques fonées sur la transformée de Fourier (TF) sont inefficaces (i.e. manque de précision) compte tenu de la particularité des données (i.e. la courte longueur). Dans cette thèse, nous proposons une nouvelle méthode pour l'estimation du vecteur de CP des signaux biomédicaux, en utilisant les informations biologiques a priori et en considérant un modèle qui représente le bruit. Les signaux enregistrés dans le cadre d'expériences développées pour le traitement du cancer ont un nombre limité de périodes. Cette information a priori peut être traduite comme la parcimonie du vecteur de CP. La méthode proposée considère l'estimation de vecteur de CP comme un problème inverse en utilisant l'inférence bayésienne générale afin de déduire toutes les inconnues de notre modèle, à savoir le vecteur de CP mais aussi les hyperparamètres (i.e. les variances associées).

**Title :** A Bayesian approach for periodic components estimation for chronobiological signals

**Keywords :** Periodic Components vector estimation, Inverse problem, Bayesian approach, Hierarchical model, Sparsity enforcing, Chronobiology

**Abstract :** The toxicity and efficacy of more than 30 anticancer agents presents very high variations, depending on the dosing time. Therefore the biologists studying the circadian rhythm require a very precise method for estimating the Periodic Components (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 seven days for the synchronization segment to two or three 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). In this thesis 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 the 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 all the unknowns of our model, i.e. the PC vector but also the hyperparameters.

