

N d'ordre : 41/2018-C/MT

**MINISTRE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE
HOUARI BOUMEDIENE
FACULTE DE MATHEMATIQUES
DEPARTEMENT DE RECHERCHE OPERATIONNELLE**



Thèse

Présentée pour l'obtention du grade de Docteur en Mathématiques
Spécialité : Recherche Opérationnelle et Mathématiques Discrètes
(ROMaD)

par

Nadia BOUSSAHA

Titre

Sur les modèles à volatilité stochastique

Soutenue publiquement, le 01/07/2018 devant le jury composé de :

Hafida GUERBYENNE,
Fayçal HAMDI,
Ourida SADKI,
Abdelouhab BIBI,
Saïd SOUAM,

Professeur à l'USTHB,
Professeur à l'USTHB,
Professeur à l'USTHB,
Professeur à l'ULBM, Oum El Bouaghi,
Professeur à l'Université de Paris Nanterre,

Présidente
Directeur de thèse
Examinatrice
Examineur
Examineur

*A mon premier enseignant de Mathématiques, à celui qui m'a inculqué la passion d'aimer les
Mathématiques, à celui qui a cru à mes choix dans la vie, à mon cher Papa...*

Remerciements

Au nom d'Allah, le Gracieux et le Miséricordieux.

En tout premier lieu, La louange est à Allah de par la grâce de qui se réalisent les bonnes choses. Je remercie le bon Dieu, le tout puissant, de m'avoir attribuée la force et la volonté pour faire aboutir ce modeste travail à son terme.

En second lieu, je tiens à exprimer, à double titre, ma sincère reconnaissance à mon directeur de thèse, Monsieur le Professeur Fayçal HAMDI, qui grâce à sa pédagogie, sa rigueur et sa passion pour les mathématiques, j'ai pu mener à bien mon projet de recherche. Par ailleurs, je le remercie de toute sa disponibilité tant bénéfique à la préparation et la finalisation de ma thèse et par-delà l'aboutissement de ce travail qui est en grande partie dû à tous les conseils que j'ai pu recevoir de sa part.

Il va de soi que j'ai à lui témoigner toute ma reconnaissance pour tous ces conseils fructueux qui ont été prodigués au cours de ces années et pour les diverses séances de travail tenues dans le cadre de ma formation et la confiance qui m'a été accordée pour travailler à ses côtés et profiter de son expérience dans différents domaines (pédagogie, administration, ...). En outre, à la fois pertinent et instructif, il a toujours su m'orienter et par conséquent la réussite de cette thèse tient essentiellement à ses directives et son expertise qui ont été d'une précieuse qualité.

En somme, grâce à son soutien constant, ses encouragements surtout durant les moments difficiles, et le fait de m'avoir introduit dans le monde des séries chronologiques, j'ai pu réaliser ma tâche d'étude et de recherche. Aucun mot ne saurait exprimer ma reconnaissance et ma gratitude à son égard, tout en lui restant infiniment redevable.

Je témoigne de toute mon estime pour le Professeur Hafida GUERBYENNE, professeur à l'USTHB, de m'avoir honorée en acceptant de présider le jury constitué à l'effet d'évaluer ma thèse. Je remercie aussi le Professeur Abdelouhab BIBI, Professeur à l'Université de Oum El Bouaghi, et le Professeur Saïd SOUAM, professeur à l'Université de Paris Nanterre, et le Professeur Ourida SADKI, professeur à l'USTHB, pour m'avoir fait l'immense honneur d'être membres de ce jury en vue d'expertiser ce travail. Tout le respect et toute la reconnaissance leur sont accordés de ma part pour l'intérêt qu'ils donnent à ce travail et tout le temps consacré à la lecture et l'étude de la thèse.

Que Monsieur Saïd SOUAM trouve ici l'expression de ma profonde gratitude pour avoir accepté de m'accueillir à maintes reprises au Laboratoire EconomiX, Université Paris Nanterre, et par-delà faciliter le déroulement de mes stages et pouvoir ainsi bénéficier de toute son aide et son soutien ainsi que des divers conseils émis, et sans oublier les nombreuses séances de travail qui m'ont été utiles pour la rédaction de ma thèse. Au-delà des travaux réalisés et liés à la thèse, j'envisage de continuer à entretenir dans le futur cette collaboration avec le Professeur Saïd SOUAM et mon directeur de thèse le Professeur Fayçal HAMDI.

Mes remerciements s'adressent également aux membres du Laboratoire RECITS, avec à leurs têtes les chefs d'équipes, Prof. Mourad BOUDHAR, Prof. Fayçal HAMDI, Prof. Hacène BELBACHIR, Prof. Miloud MIHOUBI et Prof. Mohamed El Amine CHERGUI pour leur disponibilité et les efforts consentis pour la réussite de notre formation doctorale ROMaD..., ainsi qu'aux membres de notre équipe STEP et plus particulièrement Abderaouf KHALFI et Billel ALIAT.

Mes remerciements vont également à Monsieur, Yacine BELARBI, Directeur du centre de recherche CREAD et toute l'équipe directionnelle pour tous leurs encouragements.

Ma gratitude est transmise à Monsieur Le Prof. Miloud MIHOUBI et à Monsieur Khaled BOUTARENE pour le bon accueil permanent durant toutes ces années au niveau de leur bureau (qui est en même temps celui de notre directeur de thèse), ainsi que les encouragements donnés.

Mes chaleureux remerciements sans mesure vont à ma chère mère et mon cher père, pour leur soutien indéfectible et inconditionnel, leurs sacrifices démesurés et leurs encouragements en permanence tout au long de mon cycle d'étude jusqu'à cette étape de doctorat. Je profite donc de cette occasion pour leur dire à quel point je leur suis reconnaissante de toute cette attention et cette affection ininterrompue, et nul mot ne suffirait pour les remercier. Mes vifs remerciements vont aussi à mes frères Omar et Imad Eddine et à ma petite sœur Hadia pour tout l'amour, leur présence permanente et la confiance dont ils m'ont fait preuve.

J'adresse également mes sincères remerciements à mes belles sœurs Hamida et Méria pour leurs encouragements, aide et souhaits pour ma réussite. Je remercie aussi ma petite nièce Sara Ikram pour le bonheur et la joie qu'elle apporte à ma vie.

Je suis également redevable à ma chère amie Nour El-Houda TELLACHE pour sa présence, son aide, ses suggestions et ses remarques judicieuses.

Je tiens à remercier particulièrement, mes chères amies au quotidien, Sabrina BOUCHOUKA, Imène BENRABIA, Ines BRAHIMI, Assia MEDJRREDINE, Ferial GUIDOUM et Safia Manar El-Islem BENOUMHANI pour leur soutien et leurs encouragements durant toutes ces années, créant ainsi un cadre de travail très agréable et tant bénéfique pour nos études et notre travail. Il va de soi que mes remerciements vont de pair aussi à toutes mes autres amies qui m'ont accompagnées lors de ce travail et j'espère qu'elles ne m'en voudraient pas de ne pas les avoir citées nommément mais encore une fois je les remercie toutes de leur soutien et de leurs encouragements.

Je remercie chaleureusement mes collègues Fares OUZZANI, abderaouf KHALFI, Abdelalim HASSOUNA (ZOURKANI) et Chouaïb BELDJOURDI pour leurs aides et conseils.

Une pensée particulière est accordée à mes collègues ingénieurs de soutien à la recherche du centre de recherche CREAD pour le soutien moral et les encouragements apportés.

En outre, mes remerciements doivent aller à tous ceux et à toutes celles enseignants du primaire à l'université qui ont eu la charge de la transmission du savoir et de la pédagogie y afférente.

Enfin, je remercie toutes les personnes qui ont contribué à ce travail d'une façon directe ou indirecte, continue ou ponctuelle.

Table des matières

Introduction Générale	3
1 Généralités sur les modèles espace d'états	8
1.1 Introduction	8
1.2 Modèles espace d'états	9
1.3 Filtre de Kalman	10
1.3.1 Présentation de l'algorithme	10
1.3.2 Interprétation graphique du filtre de Kalman	12
1.4 Filtre particulière	13
1.4.1 Echantillonnage d'importance	15
1.4.2 Echantillonnage d'importance séquentiel	16
1.4.3 Problème de dégénérescence	17
1.4.4 Filtre Bootstrap	21
1.4.5 Lissage particulière	23
1.4.6 Filtre particulière Rao-Blackwellisé	24
2 Modèles à Volatilité Stochastique Univariés à Coefficients Périodiques	25
2.1 Introduction	25
2.2 Modèle à Volatilité Stochastique AutoRégressif Périodique	27
2.3 Étude Probabiliste du modèle <i>PAR-SV</i>	28
2.3.1 Stationnarité périodique	28
2.3.2 Moments et structure dynamique d'un modèle <i>PAR-SV</i>	31
2.4 Estimation des paramètres d'un modèle <i>PAR-SV</i>	37
2.4.1 Méthode du quasi-maximum de vraisemblance basée sur le filtre de Kalman périodique	37

2.4.2	Méthode du maximum de vraisemblance basée sur l'algorithme <i>EM</i> et le filtre particulière	39
2.5	Application	43
2.5.1	Étude de simulation	43
2.5.2	Application empirique	49
3	Modèles à Volatilité Stochastique Multivariés à Coefficients Périodiques	56
3.1	Introduction	56
3.2	Modèle à volatilité stochastique autorégressif multivarié à coefficients périodiques	57
3.3	Étude Probabiliste du modèle <i>PVAR-SV</i>	58
3.3.1	Stationnarité périodique et calcul des moments	58
3.3.2	Fonction d'autocorrélation du modèle <i>PVAR-SV</i>	62
3.4	Estimation des paramètres d'un modèle <i>PVAR-SV</i>	65
3.4.1	Méthode du quasi-maximum de vraisemblance basée sur le filtre de Kalman périodique	66
3.4.2	Méthode du maximum de vraisemblance basée sur l'algorithme <i>EM</i> et le filtre particulière	68
3.5	Application	73
3.5.1	Données et analyse préliminaire	73
3.5.2	Résultats empiriques	77
4	Mélange de Modèles à Volatilité Stochastique	85
4.1	Introduction	85
4.2	Mélange de modèles gaussiens	86
4.2.1	Mélange de lois de probabilité	86
4.2.2	Quelques mélanges de modèles gaussiens	89
4.3	Mélange de Modèles à Volatilité Stochastique	90
4.4	Stationnarité et existence des moments	92
4.5	Estimation des paramètres d'un modèle <i>M-SV</i>	93
4.5.1	Estimation des paramètres par la méthode <i>ML</i>	93
4.5.2	Exemple illustratif en utilisant des données simulées	97
	Conclusion générale et perspectives	99

Introduction Générale

Il est généralement reconnu que la variation instantanée (ou la volatilité) des rendements de nombreuses séries financières n'est pas constante au fil du temps. La modélisation de cette volatilité joue un rôle important dans l'analyse des séries chronologiques financières. De plus, elle est depuis longtemps un sujet d'intérêt en économétrie. L'un des effets les plus connus induits par la variance des produits financiers est le phénomène de regroupement de volatilité. Celui-ci se traduit par le fait que la variance des rendements présente une autocorrélation sérielle élevée. En pratique, ceci apparaît de manière évidente sur les marchés financiers : des périodes de forte volatilité, où de grandes fluctuations dans les rendements d'un actif se suivent, et des périodes de faible volatilité où les petites fluctuations se suivent également. La modélisation de ce phénomène de volatilité, empiriquement constatée et bien mise en évidence par Mandelbrot (1963) est, pour son importance primordiale dans les études financières, un thème de recherche toujours d'actualité. De nombreux modèles captant cet effet ont été introduits et ont fait l'objet de nombreuses améliorations durant ces deux dernières décennies.

Depuis l'introduction du modèle pionnier AutoRégressif Conditionnellement Hétéroscédastique (*ARCH*) par Engle (1982), une partie importante de la littérature a été consacrée à ce phénomène. Cependant et malgré son succès, cette modélisation présente quelques limites. En effet, Nelson (1991) en relève deux. Premièrement, les contraintes imposées aux paramètres de ces modèles, pour assurer la positivité de la variance conditionnelle, sont souvent violées au cours de la procédure d'estimation. Deuxièmement, un comportement aléatoire oscillatoire du processus de variance conditionnelle est exclu.

Une autre catégorie de modèles statistiques, à savoir la classe des modèles à volatilité stochastique (*SV*) introduite par Taylor (1982), est considérée comme une alternative satisfaisante aux modèles de type *ARCH*. Cette nouvelle approche est plus flexible et capte mieux les variations dans la volatilité des actifs en permettant à la variance des rendements d'être un processus aléatoire non observé, contrairement aux modèles *ARCH* et leurs généralisations *GARCH*, où la volatilité est une fonction des observations retardées et/ou des volatilités passées. En d'autres termes, ces modèles supposent que la variance future est connue de façon certaine ce qui n'est pas tout-à-fait cohérent avec la réalité. La modélisation *SV* est motivée par la conviction que les changements dans la variance sont causés par des conditions économiques et politiques, et la volatilité ne peut pas être efficacement prédite uniquement à partir des rendements passés. La version la plus populaire du modèle *SV* considère que le logarithme de la volatilité suit un processus autorégressif de premier ordre. Notons que ceci est une approximation en temps discret du processus de diffusion d'Ornstein-Uhlenbeck en temps continu utilisé dans la littérature sur la tarification des options. De plus, le modèle *SV* permet à la log-volatilité d'évoluer au

cours du temps, ce qui assure la positivité de la variance du processus sans avoir besoin de contraintes supplémentaires, comme c'est le cas dans les modèles *ARCH/GARCH*.

Il est important de noter que les modèles *ARCH/GARCH* d'un côté, et *SV* de l'autre côté, ont des fondements théoriques différents. De ce fait, une comparaison globale de leurs performances n'est guère possible. Toutefois, de nombreuses études comparatives ont été réalisées. On peut citer à titre d'exemples Pulgarin (2001), Shephard, (1996), Kim et al. (1998), Taylor (1994), Ghysels et al. (1996) et Daniélsson (1998). Leurs résultats tendent à montrer qu'empiriquement, les modèles *SV* sont plus performants que les modèles *GARCH*(1, 1). Par ailleurs, Carnero et al. (2004) ont fait une comparaison basée sur la relation entre la Kurtosis, la persistance dans la volatilité, et la fonction d'autocorrélation de retard un. Ils ont trouvé que le modèle *SV* est plus flexible que le *GARCH*(1, 1).

Outre le regroupement de la volatilité, il existe d'autres faits stylisés importants associés aux séries des rendements financiers. En effet, Mandelbrot (1963) et Fama (1965) ont mis en évidence un ensemble de telles caractéristiques. Il est maintenant bien connu que les rendements d'un actif présentent une grande Kurtosis et ont des queues plus lourdes que la distribution normale. En outre, nous observons souvent l'effet dit de levier. Cela provoque une variance conditionnelle pour répondre de manière asymétrique aux hausses et aux chutes du prix de l'actif; les baisses d'un rendement tendent à engendrer une augmentation de la volatilité supérieure à celle induite par une hausse du rendement de même ampleur. Ces régularités empiriques, vérifiées et complétées depuis par de nombreux auteurs, apparaissent plus ou moins nettement en fonction de la fréquence d'observation de la série et de sa nature. Ces propriétés valent surtout pour des séries quotidiennes de prix d'actions.

De nombreuses extensions du modèle *SV* ont été proposées dans la littérature économétrique pour capturer différentes caractéristiques additionnelles des séries temporelles telles que la persistance à long terme des autocorrélations (mémoire longue), la dépendance sérielle et le changement de régime, l'excès de Kurtosis (voir par exemple Harvey et al. 1994; Ghysels et al., 1996; Breidt et al., 1998; Kim et Stoffer, 2008; Ghosh et al., 2015; Asai et McAleer, 2011; Asai et al., 2012). Cependant, la plupart des formulations proposées traitent les paramètres du modèle *SV* comme des constantes par rapport au temps et ne peuvent pas expliquer adéquatement les séries chronologiques dont la volatilité exhibe une certaine périodicité dans la fonction d'autocorrélation et qui ne peut pas être expliquée par les modèles *SV* avec paramètres invariants dans le temps. Il est, de ce fait, possible de considérer, similairement aux modèles *GARCH* périodiques (Bollerslev et Ghysels, 1996), une formulation à coefficients périodiques dans le temps pour l'équation de la log-volatilité.

Le modèle à Volatilité Stochastique AutoRégressive Périodique univarié (*PAR-SV*), introduit par, Aknouche et al. (2007) fournit une alternative réussie à la classe des modèles *PGARCH* en tenant compte de la variabilité et de la persistance de la volatilité, ainsi que de la caractéristique de périodicité dans la structure d'autocorrélation présente dans de nombreuses séries chronologiques.

Un autre fait stylisé, qui ne peut pas être capturé par une description univariée, est l'effet de covariance. Les volatilités des différentes séries sont souvent perçues évoluant de concert : de grands changements dans un actif sont contemporains de mouvements importants dans un autre. Cela résulte soit d'un lien formel entre un ensemble d'actifs, par exemple les taux de change, soit en raison du fait

que les actifs sont soumis au même environnement global et, par conséquent, ont des facteurs communs qui influent sur leur comportement.

Par ailleurs, grâce à l’adaptabilité des modèles à volatilité multivariés pour modéliser les variations et les corrélations dans les séries temporelles et en économétrie financière particulièrement, la littérature sur ces modèles s’est développée de manière significative. Dans ce contexte, un certain nombre de travaux discutent les modèles *GARCH* multivariés (voir par exemple Bauwens et al., 2006a ; Silvennoinen and Teräsvirta, 2009) et les modèles *SV* multivariés (voir par exemple Harvey et al., 1994 ; Daniélsson, 1998 ; Jungbacker et Koopman, 2006 ; Smith et Pitts, 2006 ; Chan et al., 2006 ; Asai et McAleer, 2009a, b). Pour une revue de littérature plus détaillée sur les modèles *SV* multivariés voir exemple Platanioti et al. (2005), Yu et Meyer (2006) et Asai et al. (2006).

Il est également observé que les distributions des résidus des modèles ajustés pour des données financières sont souvent multimodales. Une nouvelle classe de modèles à volatilité plus flexible et plus adéquate à décrire, éventuellement, plusieurs effets à la fois est la classe de mélange de modèles de séries chronologiques. Wong et Li (2001) ont proposé d’ajouter plus de flexibilité à la famille de mélange de modèles Autorégressifs *MAR* (Wong et Li, 2000), en introduisant un mélange de modèles autorégressifs à erreurs *ARCH* (*MAR-ARCH*). Plusieurs extensions du mélange de modèles *MARCH* (Wong et Li, 2001) ont été proposées dans la littérature des séries chronologiques. Citons, par exemple, le mélange de modèles *GARCH* (Haas et al., 2012 et Zhang et al., 2006), le mélange de modèle *GARCH* multivarié (Bauwens et al., 2006b), mélange de modèles *GARCH* à seuil (Giannikis et al., 2008), le mélange de modèles *ARCH* périodiques (Bentarzi et Hamdi, 2008a), le mélange de modèles *AR-ARCH* périodiques (Bentarzi et Hamdi, 2008b), le mélange de modèles *GARCH* périodiques (Hamdi et Souam, 2013, 2017) et bien d’autres.

Cependant, en comparaison avec la littérature relative aux mélanges de modèles type-*ARCH*, la littérature sur les mélanges *SV* est trop maigre. Citons, par exemple, les travaux de Chen et al. (2008), Elliott et al. (2011) et Mao et al. (2017) sur les modèles *SV* à seuil. Nous citons également les travaux de So et al. (1998) et Carvalho (2007) sur les modèles *SV* à changement de régimes Markovien. D’autres travaux ont été consacrés au mélange *i.i.d.* de modèles *SV* (Kim et al., 1998 ; Omori et al., 2007 ; Kim et Stoffer, 2008 ; et Xu et Knight, 2013).

Bien que les modèles *SV* soient considérés comme une bonne alternative aux modèles *ARCH/GARCH*, leur utilisation dans les applications empiriques est relativement limitée. Cela peut être principalement attribué aux difficultés liées à l’évaluation directe de la fonction de vraisemblance (même sous l’hypothèse de normalité), à cause du critère latent de la volatilité, ce qui présente toutefois un défi majeur quant à l’estimation des paramètres. Dit autrement, il n’existe pas une expression analytique de la densité prédictive à une étape pour les modèles *SV*. Pour cette raison, des approximations doivent être faites ou des méthodes numériques doivent être utilisées pour évaluer la fonction de vraisemblance d’un modèle *SV*¹.

Pour pallier ce problème, des progrès considérables ont été accomplis dans ce domaine, ces dernières

¹Un aperçu des différentes approches d’estimation des modèles *SV* est mis en relief dans Broto et Ruiz (2004) et Yu (2012).

années. Quatre alternatives sont connues. La première est l'estimation par la méthode des moments (*MM*) ou la méthode des moments généralisés (*GMM*) qui consiste à donner les paramètres du modèle en fonction des différents moments de la variable aléatoire étudiée (voir par exemple Taylor, 1986 ; Wiggins, 1987 ; Andersen, 1994 ; Andersen et Sørensen, 1996). La seconde est l'estimation par la méthode du quasi-maximum de vraisemblance (*QML*) et l'utilisation du filtre de Kalman sur une version linéarisée du modèle (voir par exemple Ruiz, 1994 ; Harvey et Shephard, 1996 ; Fridman et Harris, 1998). La troisième est l'estimation bayésienne ou les méthodes de Monte Carlo par Chaîne de Markov *MCMC*, qui consiste à générer un échantillon pour chaque paramètre à partir de la loi a posteriori (voir par exemple Jacquier et al., 1994 ; Kim et al., 1998 ; Chib et al., 2002). La dernière méthode est basée sur les techniques qui tentent d'évaluer la fonction de vraisemblance complète (voir par exemple Kim et Stoffer 2008 ; Chib et al., 2006 ; Carvalho et Lopes, 2007 ; Asai et McAleer, 2009a). Nous nous contenterons d'adapter, dans cette thèse, la quatrième méthode d'estimation avec la deuxième qui apparaîtront, dans notre étude, simples à adopter pour les classes de modèles proposées.

L'objectif de cette thèse est d'étudier trois généralisations du modèle *SV* classique et de montrer que l'estimation des paramètres de ces modèles peut être aisément effectuée par la méthode du maximum de vraisemblance via les algorithmes de filtrage et de lissage particulière. Nous nous intéresserons, dans un premier temps, au modèle introduit dans un document non publié par Aknouche et al. (2007), la classe de modèles à volatilité stochastique à coefficients périodiques, introduite, pour modéliser, à la fois, la variabilité instantanée et la périodicité à travers une meilleure formulation. Cette classe fournit une alternative prometteuse à la classe des modèles *PGARCH*. Dans un deuxième temps, nous développons une extension du modèle *SV* périodique au cas multivarié, pouvant capturer la périodicité dans la variance conditionnelle stochastique en utilisant une spécification qui inclut non seulement la structure de dépendance de chaque composante de la volatilité, mais aussi l'interdépendance entre différentes composantes (la causalité de Granger dans la volatilité). Dans un troisième et dernier temps, et en vue de bien décrire la multimodalité des distributions marginales et les changements récurrents de régime, nous proposons une nouvelle classe de modèles *SV*. Il s'agit de la classe du mélange de modèles à volatilité stochastique.

Pour les trois classes de modèles *SV*, nous proposons des méthodes d'estimation et nous illustrons notre proposition par des applications sur des données simulées et sur des données financières réelles. Une discussion des résultats de cette partie empirique confirme l'intérêt des modèles proposés et la performance des méthodes d'estimation adoptées.

Cette thèse est scindée en quatre chapitres. Le premier est un chapitre introductif qui porte sur la présentation du modèle espace d'états. Il contient un bref aperçu sur deux méthodes d'estimation des variables d'états : le filtrage de Kalman et le filtrage-lissage particulière. Le deuxième et le troisième chapitres décrivent le modèle à volatilité stochastique à coefficients périodiques dans sa forme univariée et sa forme multivariée respectivement. Dans chacun des deux chapitres, nous commençons par la définition du modèle. Ensuite, nous dérivons une condition sous laquelle notre modèle admet une solution périodiquement strictement stationnaire et nous obtenons, sous cette condition, l'expression explicite des moments d'ordres supérieurs. Nous étudions également leur structure d'autocovariance. Le problème de l'estimation des paramètres est traité en utilisant deux méthodes à savoir la méthode du quasi-maximum

du vraisemblance basée sur le filtre de Kalman périodique et la méthode de maximum de vraisemblance basée sur l'algorithme d'Expectation-Maximisation (EM) combiné avec le filtrage-lissage particulière. Le quatrième chapitre est consacré à la description d'un mélange de modèles à volatilité stochastique et adapte la technique fondée sur l'algorithme de filtrage particulière au problème de l'estimation. Enfin, en conclusion on peut dresser un bilan des classes de modèles proposées dans cette thèse, ainsi que les nombreux nouveaux défis et les perspectives des méthodes d'estimation pour ces modèles en particulier et les modèles à volatilité stochastique en général.

Chapitre 1

Généralités sur les modèles espace d'états

1.1 Introduction

La classe des modèles espace d'états (State Space Models) représente un cadre très puissant pour l'analyse des systèmes dynamiques. Elle repose sur la dynamique des variables d'états (inobservées) et leurs liens avec celles observées pour tirer des conclusions statistiques sur les états inobservés. La théorie des modèles espace d'états a connu des progrès considérables dans les années 1960 et a été utilisée par les économistes et les statisticiens. Elle offre notamment des outils avantageux pour analyser des séries chronologiques dans une grande variété de domaines scientifiques (Anderson et Moore, 1979 ; Harvey, 1989 ; Durbin et Koopman, 2001 ; Aoki, 2013). En effet, plusieurs modèles de séries chronologiques peuvent être exprimés sous forme espace d'états, y compris les modèles à volatilité stochastique. Ceci permet une meilleure étude et une meilleure compréhension de la dynamique de ces modèles. Les modèles espace d'états peuvent être également utilisés afin de réduire la complexité des problèmes liés à l'analyse de certains modèles de séries chronologiques.

La représentation des modèles de séries chronologiques sous forme espace d'états, a reçu une attention considérable au cours des dernières années. En effet, plusieurs travaux d'analyse des modèles linéaires stationnaires et même périodiques, à travers leurs représentations espace d'états équivalentes, ont constitué, jusqu'à présent, le centre d'intérêt de plusieurs chercheurs (voir par exemple Stoffer et Wall, 1991 ; Cavanaugh et Shumway, 1996 ; Klein et al., 2000 ; Wall et Stoffer, 2002 ; Aknouche et Hamdi, 2009 ; Hamdi, 2008 ; Aoki, 2013 ; Guerbyenne et Hamdi, 2011 ; et les nombreuses références citées).

L'objectif de ce chapitre est de donner un bref aperçu sur les modèles espace d'états. Nous donnons la définition mathématique d'un modèle espace d'états dans la Section 1.2. Ensuite, nous présentons, dans les deux dernières sections, deux méthodes pour l'estimation des variables d'états : l'algorithme de Kalman et le filtrage et le lissage particulière. Toutefois, la littérature sur le sujet étant très dense et évoluant très vite, il est, en pratique, impossible d'en donner une cartographie exhaustive. Pour cela, nous exposons les éléments de base et les grandes lignes nécessaires à la compréhension de ces procédures.

1.2 Modèles espace d'états

Les modèles espace d'états sont généralement représentés par les deux équations suivantes

$$x_t = F_t(x_{t-1}, \varepsilon_t) \quad (1.1)$$

$$y_t = H_t(x_t, e_t) \quad (1.2)$$

où F_t et H_t sont des fonctions dont la forme est connue et dépendent d'un vecteur de paramètre θ . Dans ce contexte, $x_t \in \mathbb{R}^r$ représente un état inconnu à l'instant t qui a son propre mécanisme, et est modélisé par une chaîne de Markov de distribution initiale $p_\theta(x_0)$ et de densité de transition $p_\theta(x_t|x_{t-1})$. La séquence $\{y_t, t \in \mathbb{Z}\}$ représente les données observées, qui sont supposées indépendantes conditionnellement à l'état x_t , i.e. $p_\theta(y_t|y_{t-1}, x_1, x_2, \dots, x_t) = p_\theta(y_t|x_t)$. Par conséquent, il est naturel de nommer (1.1) l'équation d'état ou d'évolution et (1.2) l'équation d'observation ou de mesure. Enfin, $e_t \in \mathbb{R}^k$, $\varepsilon_t \in \mathbb{R}^h$, et $(\varepsilon'_t, e'_t)'$ est un processus bruit blanc tel que

$$\text{cov} \left(\begin{pmatrix} \varepsilon'_t & e'_t \end{pmatrix}, \begin{pmatrix} \varepsilon'_t & e'_t \end{pmatrix} \right) = \begin{pmatrix} Q_t & S_t \\ S'_t & R_t \end{pmatrix},$$

où Q_t est une matrice définie positive et R_t est une matrice strictement définie positive. Ici, ε_t est appelé bruit (ou erreur) d'état, et e_t est appelé bruit (ou erreur) d'observation.

Une représentation graphique de ce modèle est illustrée par la Figure 1.1

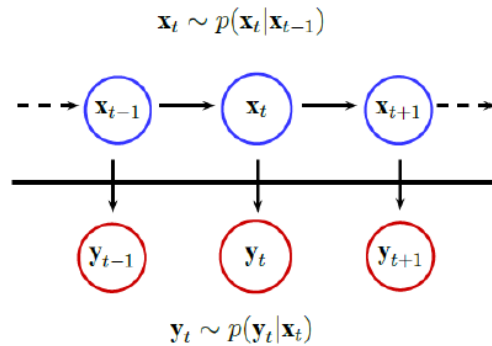


Figure 1.1. Représentation graphique du modèle espace d'états

Lorsque l'équation d'état et l'équation d'observation sont comme suit

$$x_t = A_t x_{t-1} + \varepsilon_t, \quad (1.3)$$

$$y_t = B_t x_t + e_t, \quad (1.4)$$

le modèle est référencé comme un modèle espace d'états linéaire. De plus, si ε_t et e_t dans (1.3) et (1.4) sont normalement distribuées, le modèle est appelé modèle espace d'états linéaire gaussien.

Dans cette thèse, nous discutons quelques variantes de modèles SV , là où on peut facilement percevoir que le modèle SV est un modèle espace d'états dont l'équation d'observation est non linéaire. Il est à noter que le modèle SV basique n'est pas un modèle espace d'états linéaire mais sans doute la linéarisation de l'équation d'observation le rend linéaire.

Le problème de l'estimation de l'état x_t consiste à trouver un estimateur à partir des observations $Y_s = \{y_1, \dots, y_s\}$. Trois cas de figure sont alors à distinguer :

1. Filtrage : estimation de x_t à partir de Y_s , où $t = s$.
2. Prédiction : estimation de x_t à partir de Y_s , où $t > s$.
3. Lissage : estimation de x_t à partir de Y_s , où $t < s$.

Chacun de ses problèmes peut être résolu séquentiellement par le filtre de Kalman ou le filtre particulière.

1.3 Filtre de Kalman

Le filtre de Kalman a été développé originairement par Rudolf Emil Kalman (1960) dans le contexte des systèmes linéaires. Cependant, étant donné la facilité de mise en œuvre de l'algorithme sur les calculateurs numériques, il est devenu largement utilisé dans une large gamme de domaines technologiques (radar, vision électronique, communication, etc.), et en particulier, les applications statistiques incluant les prévisions des chroniques. Le filtre de Kalman est une approche statistique, d'assimilation de données, dont le principe est de corriger la trajectoire du modèle en combinant les observations avec l'information fournie par le modèle de façon à minimiser l'erreur entre l'état vrai et l'état filtré. C'est un estimateur-prédicteur très puissant, applicable sur des modèles univariés ou multivariés. Il est également décrit comme une procédure récursive qui calcule, à un instant t dans le cas linéaire gaussien, l'estimateur optimal¹ de la variable d'état d'une façon séquentielle, en se basant sur la dynamique du modèle qui définit son évolution dans le temps, et sur l'information disponible à l'instant t . Cette dernière comprend les observations jusqu'à l'instant t et y compris le y_t . D'autre part, Schweppe (1965) a été le premier à remarquer que la fonction de vraisemblance d'un modèle écrit sous forme espace d'états, peut être évaluée au moyen du filtre de Kalman.

1.3.1 Présentation de l'algorithme

Nous nous intéressons, dans ce qui suit, au problème de filtrage et de prédiction à un pas.

Considérons maintenant le modèle espace d'états linéaire, défini par (1.3) et (1.4), tels que

$$\begin{cases} \mathbb{E}(e_t) = \mathbb{E}(\varepsilon_t) = 0, \\ \mathbb{E}(\varepsilon_t \varepsilon'_{t+h}) = \delta_{h,0} Q_t, \\ \mathbb{E}(e_t e'_{t+h}) = \delta_{h,0} R_t, \\ \mathbb{E}(e_t x'_{t-k}) = 0, \end{cases} \quad \forall t, h \in \mathbb{Z}, \forall k \geq 0,$$

où δ désigne la fonction de Kronecker. En supposant que les matrices A_t , B_t , Q_t et R_t sont connues, la prévision optimale de x_t , sur la base de l'information disponible jusqu'à l'instant $t - 1$, est dérivée d'un résultat bien connu pour des variables normales (voir par exemple Hamilton, 1994, page 92). Pour appliquer ce résultat, supposons que

¹L'optimalité ici est au sens de la minimisation des carrés des erreurs réalisées sur le vecteur d'état et au sens du maximum du vraisemblance du vecteur d'état conditionnellement au vecteur de mesure, dans ce cas les deux estimateurs sont équivalents.

- H1 Le processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ est un bruit blanc gaussien de matrice de variance covariance Q_t .
- H2 Le processus $\{e_t, t \in \mathbb{Z}\}$ est un bruit blanc gaussien de matrice de variance covariance R_t .
- H3 Le vecteur d'état initial x_0 est gaussien de moyenne μ_0 et de matrice de variance covariance P_0 .
- H4 Les bruits $\{e_t\}$ et $\{\varepsilon_t\}$, et l'état initial x_0 sont mutuellement indépendants.

L'idée du filtre de Kalman est de chercher un estimateur linéaire $\hat{x}_{t|t}$ de x_t qui vise à minimiser à chaque instant la variance de l'erreur d'estimation ; $\mathbb{E} \left[(\hat{x}_{t|t} - x_t) (\hat{x}_{t|t} - x_t)' \right]$, cet estimateur est défini par

$$\hat{x}_{t|t} = \mathbb{E} [x_t | Y_t].$$

Une propriété de cet estimateur est son écriture récursive qui définit les équations du filtre de Kalman discret. Cette écriture permet également de déterminer récursivement la variance de l'erreur de filtrage.

Le filtre de Kalman est structuré en deux étapes reprises d'itération en itération. Les deux premières équations (1.5) et (1.6) sont des équations de «mise à jour du temps» (étape de prédiction) et les deux suivantes (1.7) et (1.8) de «mises à jour des mesures» (étape de correction). La première étape concerne les lois de probabilité *a priori*. Cette dernière ne dépend pas des observations à l'instant t et le calcul peut être fait «hors-ligne», i.e. sans tenir compte de la variable observée courante y_t . La seconde étape, différemment de la première, concerne les lois de probabilité *a posteriori* qui prennent en compte l'information à l'instant t . Enfin, la pondération K_t , appelée matrice de gain, qui intervient dans les équations précédentes, est actualisée à chaque itération par la dernière équation (1.9). Chaque itération de l'algorithme se résume par les cinq équations suivantes

$$\hat{x}_{t|t-1} = A_t \hat{x}_{t-1|t-1}, \tag{1.5}$$

$$P_{t|t-1} = A_t P_{t-1|t-1} A_t' + Q_t, \tag{1.6}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - B_t \hat{x}_{t|t-1}), \tag{1.7}$$

$$P_{t|t} = (I - K_t B_t) P_{t|t-1}, \tag{1.8}$$

$$K_t = A_t P_{t|t-1} B_t' (B_t P_{t|t-1} B_t' + R_t)^{-1}, \tag{1.9}$$

et par l'initialisation de l'état du système et de sa matrice de covariance

$$\hat{x}_{1|0} = x_0 \text{ et } \hat{P}_{1|0} = P_0.$$

Ici, $\hat{x}_{t|t-h}$, $h = 0, 1$, est le meilleur prédicteur linéaire du vecteur d'état x_t à l'instant t à partir des mesures disponibles jusqu'à l'instant $t - h$, et $P_{t|t-h} = \mathbb{E} \left[(x_t - \hat{x}_{t|t-h}) (x_t - \hat{x}_{t|t-h})' \right]$ est la matrice de covariance de l'erreur de prévision correspondante.

L'équation (1.5) permet la prédiction de l'évolution temporelle à un pas de l'état x_t . Cette prédiction est notée par $\hat{x}_{t|t-1}$ et elle est définie comme la projection de x_t sur son passé (passé synthétisé par x_{t-1}). L'équation (1.7) réactualise l'estimation de l'état x_t . Cette réactualisation est notée par $\hat{x}_{t|t}$ et elle est comme la somme pondérée de la prévision $\hat{x}_{t|t-1}$ et un terme correctif proportionnel à l'erreur relative à la dernière observation y_t .

L'équation (1.9) calcule la matrice de gain optimal K_t , qui tient compte des caractéristiques statistiques du bruit de mesure mais ne dépend pas des données mesurées. Donc elle peut être calculée a priori.

Les équations (1.6) et (1.8) sur les matrices de covariance sont appelées «équations de Riccati». Ces équations permettent de calculer la suite des gains de Kalman K_t . La matrice de covariance a posteriori $P_{t|t}$ connaît généralement un gain en précision par rapport à la matrice de covariance a priori $P_{t|t-1}$ grâce au terme $K_t B_t P_{t|t-1}$. La matrice de covariance a priori en t , $P_{t|t-1}$, tient compte des erreurs liées aux innovations de l'état avec la matrice Q_t , mais est aussi augmentée d'un terme $A_t P_{t|t} A_t'$ associé aux erreurs sur l'état à la date t .

1.3.2 Interprétation graphique du filtre de Kalman

Une interprétation du filtre de Kalman repose sur l'utilisation des propriétés élémentaires de l'algèbre linéaire. Rechercher $\hat{x}_{t|t} = \mathbb{E}[x_t | Y_t]$, revient en effet à projeter x_t sur le sous-espace engendré par $\{y_1, y_2, \dots, y_t\}$. Or cet espace peut se décomposer en l'espace des valeurs observées par le passé (engendré par $\{y_1, y_2, \dots, y_{t-1}\}$) et celui de l'innovation \hat{e}_t (égale à $y_t - B_t \hat{x}_{t|t-1}$). En projetant x_t sur ces deux sous-espaces, on obtient, par définition, $\hat{x}_{t|t-1}$ et $\mathbb{E}[x_t | \hat{e}_t]$. On peut alors montrer que $\mathbb{E}[x_t | \hat{e}_t]$ peut se réécrire comme $K_t \hat{e}_t$, où K_t est la matrice de gain calculée dans l'équation (1.9). L'équation (1.7) exprime alors la projection recherchée comme la somme des projections sur chaque sous-espace (voir Figure 1.2).

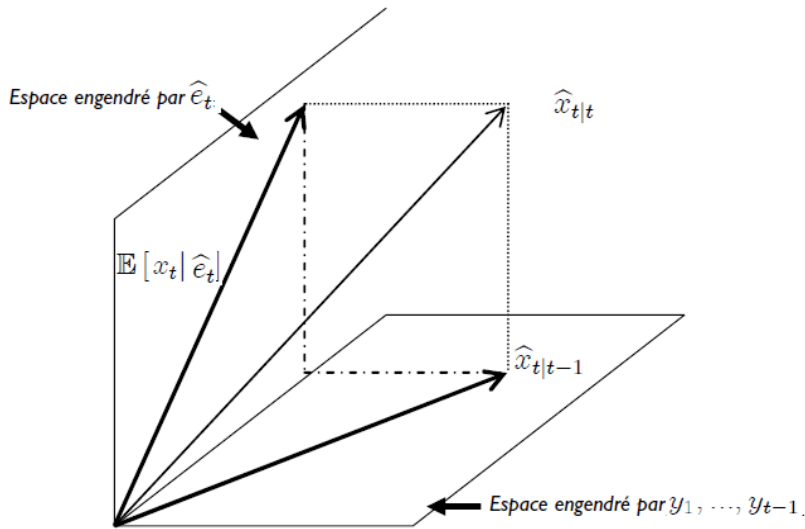


Figure 1.2. Projection du vecteur \hat{x}_t sur le sous-espace engendré par $\{y_1, y_2, \dots, y_t\}$.

Les équations (1.5)-(1.9) expriment les grandeurs a priori, en fonction de celles a posteriori et réciproquement. Il est possible de condenser ces deux étapes en une seule, conduisant ainsi aux formes prédicteur et correcteur du filtre de Kalman suivant.

Algorithme 1.1 (Filtre de Kalman, 1960) ¹

$$\begin{aligned}\Omega_t &= B_t P_{t|t-1} B_t' + R_t, \\ K_t &= A_t P_{t|t-1} B_t' \Omega_t^{-1}, \\ \hat{x}_{t+1|t} &= (A_{t+1} - K_t B_t) \hat{x}_{t|t-1} + K_t y_t, \\ P_{t+1|t} &= A_{t+1} (P_{t|t-1} - P_{t|t-1} \Omega_t^{-1} P_{t|t-1}) A_{t+1}' + Q_{t+1}.\end{aligned}$$

Le filtre de Kalman, comme mentionné précédemment, apporte une solution optimale dans le cas des modèles linéaires gaussiens. En revanche, lorsque le modèle est non-linéaire et/ou non-gaussien, le filtre de Kalman reste uniquement optimal parmi les estimateurs linéaires. Dans le cas où le modèle est non gaussien, on peut approximer la densité a posteriori par une distribution gaussienne. Dans le cas où le modèle est non-linéaire, le filtre de Kalman étendu (Extended Kalman Filter, *EKF*) nécessite la linéarisation locale des fonctions non-linéaires à l'aide d'un développement en série de Taylor tronqué à l'ordre 1 au voisinage de la variable estimée $\hat{x}_{t|t}$ courante et l'application du filtre de Kalman classique.

Toutefois, ces techniques ont parfois des performances peu satisfaisantes, en particulier dans les modèles dont la densité a posteriori présente un caractère très pointu et/ou multimodal, et son approximation par une distribution gaussienne n'est plus envisageable et peut entraîner des erreurs d'estimation importantes. Dans ce cas, il faut avoir recours à des méthodes d'approximation numériques comme les méthodes de Monte Carlo, afin d'approcher n'importe quel type de densité a posteriori.

1.4 Filtre particulaire

Les filtres particuliers sont des méthodes séquentielles de Monte Carlo qui peuvent être appliqués aux modèles espace d'états. Ils peuvent être considérés comme une alternative puissante au filtre de Kalman classique pour les problèmes d'estimation optimale d'un modèle non-linéaire et/ou non-gaussien. En raison de la popularité des méthodes de particules, quelques surveys ont déjà été publiés sur le sujet (Arulampalam et al., 2002; Cappé et al., 2007; Fearnhead, 2008; Künsch, 2001; Doucet et Johansen, 2012).

Contrairement au filtre de Kalman qui fournit une seule estimation $\hat{x}_{t|s} = \mathbb{E}[x_t|Y_s]$, l'idée principale des filtres particuliers est d'approximer la densité de probabilité *a posteriori* $p(x_t|Y_s)$ de la variable d'état x_t (ou $p(X_t|Y_s)$ de la trajectoire $X_t = (x_0, \dots, x_t)$) à l'instant t conditionnellement aux mesures Y_s , par une distribution empirique, i.e. en utilisant un échantillon discret obtenu par simulation.

Il est bien connu que l'estimation bayésienne nécessite le calcul des intégrales du type

$$\mathbf{I}(h) = \int h(x_t) p(x_t|Y_t) dx_t. \quad (1.10)$$

Dans le cas de densités de probabilité relativement complexes, le calcul analytique de cette intégrale n'est pas facile et les techniques classiques d'intégration ne sont pas suffisamment précises et/ou sont trop coûteuses en terme de temps de calcul.

¹Pour la preuve voir par exemple Brockwell et Davis (2016).

Considérons $\{x_t^{(j)}\}_{j=\overline{1,M}}$ une suite de réalisations *i.i.d.* à partir de $p(x_t|Y_t)$, appelées particules, i.e. $x_t^{(j)} \sim p(\cdot|Y_t)$, $j = \overline{1,M}$. Une approximation de Monte Carlo de la densité a posteriori est donnée par la distribution empirique

$$P_M(dx_t|Y_t) = \frac{1}{M} \sum_{j=1}^M \delta_{x_t^{(j)}}(dx_t).$$

où $\delta_x(\cdot)$ est la fonction de Dirac au point x . En se basant sur cette approximation, $\mathbf{I}(h)$ peut être approchée par

$$\widehat{\mathbf{I}}_M(h) = \int h(x_t) P_M(dx_t|Y_t) = \frac{1}{M} \sum_{j=1}^M h(x_t^{(j)}). \quad (1.11)$$

Il est possible de montrer que cette estimation de l'intégrale $\mathbf{I}(h)$ est non biaisée. De plus, si la variance a posteriori de $h(x_t)$, notée σ_h^2 , telle que

$$\sigma_h^2 \triangleq \mathbb{E}_{p(x_t|Y_t)} [h^2(X_t)] - \mathbf{I}^2(h),$$

est finie, alors la variance de $\widehat{\mathbf{I}}_M(h)$ est donnée par

$$\text{var}(\widehat{\mathbf{I}}_M(h)) = \frac{\sigma_h^2}{M}.$$

Ainsi, d'après la loi forte des grands nombres, $\widehat{\mathbf{I}}_M(h)$ converge presque sûrement vers $\mathbf{I}(h)$, lorsque M tend vers l'infini,

$$\widehat{\mathbf{I}}_M(h) \xrightarrow[M \rightarrow \infty]{p.s.} \mathbf{I}(h).$$

et la vitesse de convergence est donnée par le théorème central limite

$$\sqrt{M} \left[\widehat{\mathbf{I}}_M(h) - \mathbf{I}(h) \right] \xrightarrow[M \rightarrow \infty]{} \mathcal{N}(0, \sigma_h^2).$$

Ce constat met en avant un point important des méthodes de Monte Carlo. En effet, à partir d'un ensemble de M particules $\{x_t^{(j)}\}_{j=\overline{1,M}}$, la relation (1.11) permet d'obtenir facilement des estimateurs de l'intégrale $\mathbf{I}(h)$ pour toute fonction h . De plus, la vitesse de convergence du calcul de cette approximation dépend uniquement de leur effectif et non pas de la dimension de l'espace d'états. Avec les méthodes classiques d'intégration, la vitesse de convergence augmente de façon exponentielle avec la dimension de l'espace d'états. Cependant, ce résultat est obtenu en supposant que l'on dispose d'un échantillon distribué selon la loi visée. Le problème consiste alors à produire cet échantillon, c'est à dire à simuler un ensemble de réalisations de la variable aléatoire distribuée selon la densité de probabilité *a posteriori* $p(x_t|Y_t)$. Malheureusement, un problème se pose en pratique; il est souvent impossible de générer des échantillons selon cette densité. Autrement dit, l'obtention des particules n'est pas un problème trivial et le recours à des algorithmes d'échantillonnage est nécessaire.

Il existe de nombreuses méthodes d'échantillonnage selon une loi $p(x_t|Y_t)$, notamment la méthode de la transformation inverse, la méthode d'acceptation/rejet, l'échantillonnage d'importance (Importance Sampling (*IS*)), les méthodes de Monte Carlo par chaînes de Markov (Markov Chain Monte Carlo (*MCMC*)), etc. Nous présentons ci-dessous l'échantillonnage d'importance.

1.4.1 Echantillonnage d'importance

L'objectif de l'échantillonnage d'importance, appelé aussi échantillonnage préférentiel, est de générer des particules dans les régions d'importance de la densité a posteriori, c'est-à-dire dans les régions de l'espace d'états dans lesquelles la densité a posteriori prend des valeurs élevées. Il repose sur l'introduction d'une densité de probabilité $q(x_t|Y_t)$ appelée densité d'importance, selon laquelle il est possible de générer des échantillons, et dont le support contient celui de $p(x_t|Y_t)$. Sous ces hypothèses, l'intégrale définie dans l'équation (1.10) peut être réécrite sous la forme

$$\begin{aligned} \mathbf{I}(h) &= \int h(x_t) p(x_t|Y_t) dx_t \\ &= \int h(x_t) \frac{p(x_t|Y_t)}{q(x_t|Y_t)} q(x_t|Y_t) dx_t \\ &= \frac{\int h(x_t) w(x_t) q(x_t|Y_t) dx_t}{\int w(x_t) q(x_t|Y_t) dx_t}, \end{aligned} \tag{1.12}$$

où $w(x_t)$ est le poids d'importance non normalisé, défini par

$$w(x_t) = w_t = \frac{p(x_t|Y_t)}{q(x_t|Y_t)}. \tag{1.13}$$

Considérons maintenant $\{x_t^{(j)}\}_{j=1, \dots, M}$ une suite de réalisations *i.i.d.* générée à partir de $q(x_t|Y_t)$. En appliquant la méthode de Monte Carlo au numérateur et au dénominateur de l'équation (1.12), on obtient une approximation de $\mathbf{I}(h)$. Cette approximation est donnée par

$$\hat{\mathbf{I}}_M(h) = \frac{\frac{1}{M} \sum_{j=1}^M w(x_t^{(j)}) h(x_t^{(j)})}{\frac{1}{M} \sum_{j=1}^M w(x_t^{(j)})} = \sum_{j=1}^M w_t^{(j)} h(x_t^{(j)}), \tag{1.14}$$

où $w_t^{(j)}$ est le poids d'importance normalisé, donnée par

$$w_t^{(j)} = \frac{w(x_t^{(j)})}{\sum_{i=1}^M w(x_t^{(i)})}.$$

Illustration graphique

Nous cherchons à approcher la densité

$$p(x) = 0.4\mathcal{N}(-2, 2) + 0.3\mathcal{N}(2, 2) + 0.3\mathcal{N}(5.5, 0.5).$$

La densité d'importance choisie est une distribution gaussienne, $q(x) = \mathcal{N}(2, 10)$. Nous considérons un ensemble de $M = 30$ particules distribuées selon la densité d'importance q . Nous visons à explorer les zones "pertinentes" de l'espace d'états. Or, par hypothèse, l'échantillon est distribué selon la loi p , donc les valeurs de x pour lesquelles $q(x)$ est supérieur à $p(x)$ sont sur-représentées et celles, au contraire, pour lesquelles $q(x)$ est inférieur à $p(x)$ sont sous-représentées (voir Figure 1.3). L'idée¹ consiste alors à

¹Dans la méthode d'acceptation/rejet, soit on accepte la particule et on l'associe au poids $1/M$, soit elle est rejetée. Par contre dans l'échantillonnage d'importance on accepte toutes les particules (aucune particule n'est rejetée), mais en pondérant les particules avec un poids normalisé $w^{(j)}$.

introduire pour chacune des valeurs de l'échantillon une pondération afin de tenir compte de son poids $w^{(j)}$ dans le calcul de l'espérance de $h(x)$.

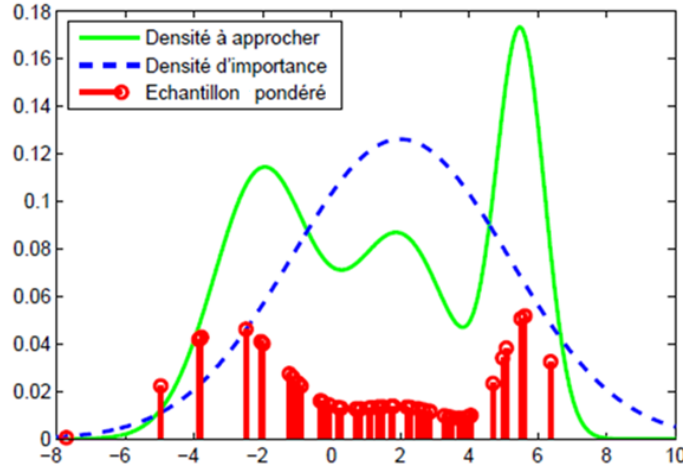


Figure 1.3. Echantillonnage d'importance.

La densité de la loi $0.4\mathcal{N}(-2, 2) + 0.3\mathcal{N}(2, 2) + 0.3\mathcal{N}(5.5, 0.5)$ est approchée avec un système de $M = 30$ particules pondérées distribuées selon la densité d'importance de la loi $\mathcal{N}(2, 10)$.

L'échantillonnage d'importance, tel qu'il est présenté ici, est inadapté pour la mise en œuvre d'une procédure séquentielle pour l'estimation de la densité a posteriori. En effet, lorsqu'une nouvelle observation y_t est disponible, il faut redéterminer l'ensemble des particules pour tous les instants. On obtient donc une complexité de calculs qui augmente avec le temps. Cependant, l'échantillonnage d'importance peut être modifié afin que le calcul d'une approximation de la densité a posteriori puisse être réalisé sans avoir besoin de modifier l'ensemble des trajectoires des particules $\{x_{t-1}^{(j)}\}_{j=1, M}$. La méthode correspondante porte le nom d'échantillonnage d'importance séquentiel (Sequential Importance Sampling (SIS)).

1.4.2 Echantillonnage d'importance séquentiel

L'aspect séquentiel du problème d'estimation de la densité a posteriori se résume dans l'utilisation du résultat à l'instant $t - 1$ pour estimer la densité à l'instant t . On rappelle que la densité a posteriori à l'instant t admet, à cause du caractère markovien de la variable x_t , la formulation récursive suivante :

$$p(X_t|Y_t) = p(X_{t-1}|Y_{t-1}) \frac{p(y_t|x_t) p(x_t|x_{t-1})}{p(y_t|Y_{t-1})}. \quad (1.15)$$

Afin d'appliquer séquentiellement l'algorithme d'échantillonnage d'importance, la densité d'importance doit vérifier la forme récursive suivante :

$$\pi_t(X_t|Y_t) = \pi_{t-1}(X_{t-1}|Y_{t-1}) q_t(x_t|X_{t-1}, Y_t) = \pi_0(x_0) \prod_{i=1}^t q_i(x_i|X_{i-1}, Y_i). \quad (1.16)$$

A l'instant $t = 0$, les particules sont échantillonnées à partir de $\pi_0(x_0)$. La densité d'importance $q_t(x_t|X_{t-1}, Y_t)$ permet de propager les particules de l'instant $t - 1$ à l'instant t . On peut noter que la propagation augmente la longueur de la trajectoire du vecteur X_t . Construites de cette manière, les particules à l'instant t sont distribuées selon $\pi_t(X_t|Y_t)$. Pour corriger cette différence avec la densité a posteriori, il est nécessaire de pondérer les particules en calculant les poids d'importance. En utilisant les équations (1.15) et (1.16) dans l'équation (1.13), le mécanisme de mise à jour des poids obéit alors à l'expression récursive suivante

$$w_{t|t}^{(j)} = w_{t-1|t-1}^{(j)} \frac{p(y_t|x_t^{(j)}) p(x_t^{(j)}|x_{t-1}^{(j)})}{q_t(x_t^{(j)}|x_{t-1}^{(j)}, Y_{t-1}) p(y_t|Y_{t-1})}$$

ou

$$w_{t|t}^{(j)} \propto w_{t-1|t-1}^{(j)} \frac{p(y_t|x_t^{(j)}) p(x_t^{(j)}|x_{t-1}^{(j)})}{q_t(x_t^{(j)}|x_{t-1}^{(j)}, Y_{t-1})}. \quad (1.17)$$

Il est à noter que l'utilisation d'une formule définie à une constante près ne pose pas de problème car les poids sont ensuite normalisés.

L'algorithme suivant récapitule la procédure de l'échantillonnage d'importance séquentiel.

Algorithme 1.2 (Echantillonnage d'importance séquentiel (SIS))

1. Pour $t = 0$, générer un échantillon $\{x_t^{(j)}\}_{j=\overline{1, M}}$ à partir de $q(x_0)$ et poser $w_0^{(j)} = \frac{1}{M}$.

2. Pour $t > 0$,

(a) Générer un échantillon $\{x_t^{(j)}\}_{j=\overline{1, M}}$ à partir de $q_t(x_t^{(j)}|x_{t-1}^{(j)}, Y_{t-1})$.

(b) Mise à jour des poids : $w_t^{(j)} \propto w_{t-1}^{(j)} \frac{p(y_t|x_t^{(j)}) p(x_t^{(j)}|x_{t-1}^{(j)})}{q_t(x_t^{(j)}|x_{t-1}^{(j)}, Y_{t-1})}$, $j = \overline{1, M}$.

(c) Normalisation des poids : $w_t^{(j)} = w_t^{(j)} / \sum_{i=1}^M w_t^{(i)}$, $j = \overline{1, M}$.

(d) Finalement, la séquence des M particules $\{x_t^{(j)}\}_{j=\overline{1, M}}$ est un échantillon aléatoire issu de $p(x_t|Y_t)$ (on dit aussi que la séquence des M trajectoires $\{X_t^{(j)}\}_{j=\overline{1, M}}$ est un échantillon aléatoire issu de $p(X_t|Y_t)$).

1.4.3 Problème de dégénérescence

Une limitation connue de l'algorithme d'échantillonnage d'importance séquentiel est que la variance des poids augmente avec le temps. Doucet et al. (2000) ont montré¹ que la plupart des poids des pondérations diminuent dans le temps. Ce phénomène est appelé la dégénérescence des pondérations et se traduit par une diminution significative du nombre de particules utiles au cours du temps. En

¹La preuve de cette proposition est une extension directe du théorème de Kong et al. (1994) dans le cas d'une densité d'importance.

effet, les poids normalisés décroissent vers 0 pour toutes les trajectoires simulées sauf une qui prend tout le poids et ceci à cause du caractère multiplicatif de la formule récursive des poids (1.17). Ce qui conduit à fausser l'approximation de la densité a posteriori et provoque à long terme la divergence du filtre, autrement dit elle entraîne des effets néfastes sur la qualité de la précision de l'estimateur (1.14). Cette dégénérescence signifie qu'un grand effort de calcul est consacré à la mise à jour des particules dont la contribution à l'approximation est presque nulle. Arulampalam et al. (2002) ont mentionné, pour remédier à ce problème, deux critères qui sont le bon choix de la densité d'importance et le rééchantillonnage.

Choix de la densité d'importance

La densité d'importance joue un rôle primordial dans les algorithmes de filtrage particulaire et son choix conditionne le bon fonctionnement du filtre. Elle doit être choisie pour minimiser la variance des poids. Les principaux choix possibles sont (voir Doucet et al., 2000)

1. *Densité d'importance optimale* introduite par Zaritskii et al. (1975) et est donnée par

$$q\left(x_t | X_{t-1}^{(j)}, Y_t\right) = p\left(x_t | x_{t-1}^{(j)}, y_t\right)$$

avec les poids associés

$$w_t^{(j)} = w_{t-1}^{(j)} \int p(y_t | x_t) \left(x_t | x_{t-1}^{(j)}, y_t\right) dx_t.$$

Ce choix est optimal au sens où cette densité minimise la variance des poids conditionnellement à l'état précédent x_{t-1} et à l'observation actuelle y_t . Cependant, ce choix est rarement applicable dans la pratique, car cette densité optimale souffre de deux inconvénients majeurs : il n'est pas toujours possible d'échantillonner selon la densité $p(x_t | x_{t-1}, y_t)$, et aussi d'évaluer cette intégrale par rapport au nouvel état. En revanche, il existe deux cas où l'utilisation de la densité d'importance optimale est possible¹. Pour de nombreux autres modèles, de telles évaluations sont impossibles. Nous présentons maintenant deux méthodes sous-optimales qui permettent l'approximation de la fonction d'importance optimale.

2. *Densité d'importance obtenue par linéarisation locale* : Doucet et al. (2000) présentent deux schémas qui aboutissent à une fonction d'importance gaussienne dont les paramètres sont évalués en utilisant des linéarisations locales, c'est-à-dire qui dépendent de la trajectoire simulée $j = \overline{1, M}$. Une telle approche semble être une manière très prometteuse de procéder avec de nombreux modèles, où les linéarisations sont facilement disponibles (voir par exemple Pitt et Shephard, 1999).

- (a) *Linéarisation locale du modèle espace d'états* : une linéarisation locale du modèle de la même manière que le filtre de Kalman étendu (*EKF*) est proposée. Cependant, dans ce cas, cette linéarisation est effectuée dans le but d'obtenir une fonction d'importance et l'algorithme obtenu converge toujours asymptotiquement vers la densité d'importance optimale. En procédant ainsi, on approche, en réalité, la densité d'importance optimale par une densité gaussienne, i.e.

$$q\left(x_t | x_{t-1}^{(j)}, y_t\right) = \mathcal{N}\left(x_{t|t}^{(j)}, \Sigma_{t|t}^{(j)}\right),$$

¹Voir Doucet et al. (2000) et Arulampalam et al. (2002).

dont la moyenne $x_{t|t}^{(j)}$ et la matrice de covariance $\Sigma_{t|t}^{(j)}$ sont calculées¹ à partir du filtre de Kalman étendu. Ce choix est efficace, mais assez coûteux à mettre en œuvre en termes de nombre d'opérations, car à chaque instant le filtre de Kalman est exécuté.

- (b) *Linéarisation locale de la fonction d'importance optimale* : on suppose ici que $l(x_t) = \log p(x_t | x_{t-1}^{(i)}, y_t)$ est deux fois différentiable en x_t . En utilisant le développement de Taylor d'ordre deux au voisinage de x_t , on suggère de prendre une loi gaussienne $\mathcal{N}(m_t^{(j)}, R_t^{(j)})$ comme loi d'importance, telle que $m_t^{(j)}$ et $R_t^{(j)}$ sont calculées² en fonction de la dérivée première et seconde de l par rapport à x_t .

3. *Loi d'évolution de l'état* : il est souvent pratique de choisir la loi d'évolution comme densité d'importance, i.e.

$$q(x_t | x_{t-1}^{(j)}, y_t) = p(x_t | x_{t-1}^j),$$

dans ce cas les poids sont calculés comme suit

$$w_t^{(j)} = w_{t-1}^{(j)} p(y_t | x_t^{(j)}).$$

L'algorithme ainsi construit est appelé *filtre Bootstrap* adopté par Gordon et al. (1993). Il a l'avantage d'être facile à implémenter. Cependant, il est souvent peu efficace vu que l'information apportée à l'instant t par la nouvelle observation n'est pas prise en compte pour la mise à jour des particules; ce qui se traduit par une grande variance des poids.

Rééchantillonnage des particules

Si le système de particules pondérées $\{x_t^{(j)}, w_t^{(j)}\}_{j=1, \overline{M}}$ est tel que les poids ont une grande variance, on va générer un autre système de particules de même taille en favorisant les particules de grands poids. Celles-ci seront dupliquées, au détriment de celles, faiblement pondérées, qui disparaissent, tout en restant à un nombre de particules constant. On obtient alors un nouvel échantillon de particules de même poids. Gordon et al. (1993) ont, les premiers, introduit une étape de rééchantillonnage dans le cadre de l'échantillonnage d'importance séquentiel. On peut assimiler le rééchantillonnage à la transformation de la mesure pondérée discrète $\{x_t^{(j)}, w_t^{(j)}\}$ en une mesure équi-pondérée discrète $\{x_t^{(j)}, 1/M\}$. Citons brièvement, les principales méthodes de sélection des particules classiquement utilisées dans les filtres bootstrap comme les résumés Arulampalam et al. (2002).

- *Le rééchantillonnage classique ou multinomial* : Cette technique est la plus utilisée. Soit $M_t^{(j)}$ le nombre de particules descendants engendrés par la $j^{\text{ème}}$ particule à l'instant t . Le but recherché est que $\mathbb{E}(M_t^{(j)}) = M w_t^{(j)}$, ce qui signifie que plus le poids d'une particule est élevé, plus son nombre de descendants espéré est grand. Ceci peut être réalisé grâce à un tirage multinomial.
- *Le rééchantillonnage résiduel* : Cette technique est une amélioration du tirage multinomial, dans le but est de diminuer la variance des particules dupliquées. L'idée dans une première étape, est

¹Pour les formules de la moyenne et la matrice de variance-covariance, ainsi que les poids associés, voir par exemple Doucet et al. (2000).

²cf. Doucet et al. (2000).

de choisir de façon déterministe une partie des particules, puis dans la deuxième étape utiliser la redistribution multinomiale pour sélectionner les particules résiduelles.

En définissant une structure générique basée sur l'algorithme *SIS*, auquel on ajoute une étape de rééchantillonnage, l'algorithme obtenu est appelé échantillonnage d'importance séquentiel avec rééchantillonnage (Sampling Importance Resampling, *SIR*).

Indicateur de dégénérescence

Notons que cette étape de redistribution ne doit pas être déclenchée à chaque cycle de calcul, car elle entraîne des erreurs de Monte Carlo : absence de diversité des particules (une particule est dupliquée plusieurs fois). L'idéal est de réaliser un compromis concernant la fréquence du rééchantillonnage en ne faisant intervenir ce dernier que lorsque cela s'avère nécessaire, en choisissant le moment propice pour le rééchantillonnage des particules avant la dispersion des poids. Cette étape doit être effectuée lorsque l'efficacité du filtre (liée au nombre de particules) se situe en dessous d'un certain seuil. Ce seuillage a pour effet de favoriser la diversité des particules en limitant le nombre d'appels à la procédure de rééchantillonnage.

Plusieurs méthodes heuristiques, dont le but est de mesurer cette dégénérescence, ont été proposées dans la littérature. On citera deux critères parmi les critères les plus utilisés. Le premier critère a été défini par Kong et al. (1994), il est basé sur la variance des pondérations. Il introduit le paramètre N_{eff}^t , appelé *nombre de particules efficaces*, inversement proportionnel à la variance de l'estimateur

$$N_{eff}^t = \frac{M}{1 + \frac{\text{var}(w_t)}{[\mathbb{E}(w_t)]^2}} = \frac{M}{1 + M^2 \text{var}(w_t)}.$$

Il vaut M lorsque les particules sont réparties uniformément, et 1 lorsque toutes les particules sauf une admettent un poids nul. Notons que l'expression analytique de ce scalaire ne peut pas être évaluée exactement. Cependant il est possible de l'estimer à tout instant t au moyen de la quantité

$$\widehat{N}_{eff}^t = \frac{1}{\sum_{j=1}^M (w_t^{(j)})^2}.$$

Il suffit alors de fixer pour N_{eff}^t un seuil (fonction du nombre total de particules M à l'instant t , noté n^T) en dessous duquel une procédure de rééchantillonnage est à envisager.

Le second critère de rééchantillonnage est basé sur le calcul de l'entropie des pondérations. L'emploi de cet indicateur, introduit par Pham (2001), se justifie par le fait qu'une analogie est possible entre les particules d'un filtre de Monte Carlo et celle d'un gaz. L'entropie Ent s'obtient comme suit

$$Ent_t = \log M + \sum_{j=1}^M w_t^{(j)} \log w_t^{(j)}.$$

Nous avons toujours $0 < Ent_t \leq \log(M)$. Ce critère est nul si tous les poids sont égaux à $1/M$ et vaut $\log(M)$ si l'un des poids vaut 1. Ainsi, de même que pour l'indicateur du nombre effectif, il suffit de fixer pour Ent_t un seuil (noté Th) en dessus duquel un rééchantillonnage devra être déclenché.

1.4.4 Filtre Bootstrap

La méthode d'estimation des paramètres des modèles SV dans cette thèse est basée sur le filtre particulière connu sous la terminologie de Bootstrap. Le filtre Bootstrap possède quelques propriétés intéressantes qui justifient l'intérêt qui lui est porté.

- Il est rapidement et facilement implantable.
- Il est aisément adaptable à un nouveau problème. Il suffit de modifier, au niveau du programme, la distribution d'importance ainsi que les poids d'importance, ce qui paraît clairement dans l'estimation des trois modèles proposés (voir Chapitres 2, 3 et 4).

L'algorithme 1.3 et la Figure 1.4 suivants exposent schématiquement les différentes phases du fonctionnement d'un filtre Bootstrap. Notons que $f_t^{(j)}$ est la $j^{\text{ème}}$ particule filtrée à l'instant t pour approximer $p(x_t|Y_t)$, et $p_t^{(j)}$ est la $j^{\text{ème}}$ particule prédite à l'instant t pour approximer $p(x_t|Y_{t-1})$.

Algorithme 1.3 (Filtre Bootstrap (Gordon et al., 1993)) ¹

1. Poser $w_0^{(j)} = \frac{1}{M}$, et générer un échantillon $\left\{f_t^{(j)}\right\}_{j=\overline{1, M}}$ à partir de $p(x_0)$.
2. calculer $p_t^{(j)} = F_t\left(f_{t-1}^{(j)}, w_{t-1}^{(j)}\right)$, $j = \overline{1, M}$.
3. Mise à jour des poids : $w_t^{(j)} \propto w_{t-1}^{(j)} p\left(y_t|x_t^{(j)}\right)$, $j = \overline{1, M}$.
4. Normalisation des poids : $w_t^{(j)} = w_{t-1}^{(j)} / \sum_{j=1}^M w_{t-1}^{(j)}$, $j = \overline{1, M}$.
5. Calcul de \widehat{N}_{eff} (ou \widehat{Ent}),
 - (a) Si $\widehat{N}_{eff} \geq n^T$ (ou $\widehat{Ent} > Th$) alors générer $\left\{f_t^{(j)}\right\}_{j=\overline{1, M}}$ par rééchantillonnage de $\left\{p_t^{(i)}, w_t^{(i)}\right\}_{i=\overline{1, M}}$ et poser $w_t^{(j)} = \frac{1}{M}$.
 - (b) Sinon $f_t^{(j)} = p_t^{(j)}$, $j = \overline{1, M}$
6. Finalement, la séquence des M particules $\left\{f_t^{(j)}\right\}_{j=\overline{1, M}}$ est un échantillon aléatoire de $p(x_t|Y_t)$.

¹Une mesure de dégénérescence est intégrée dans cet algorithme.

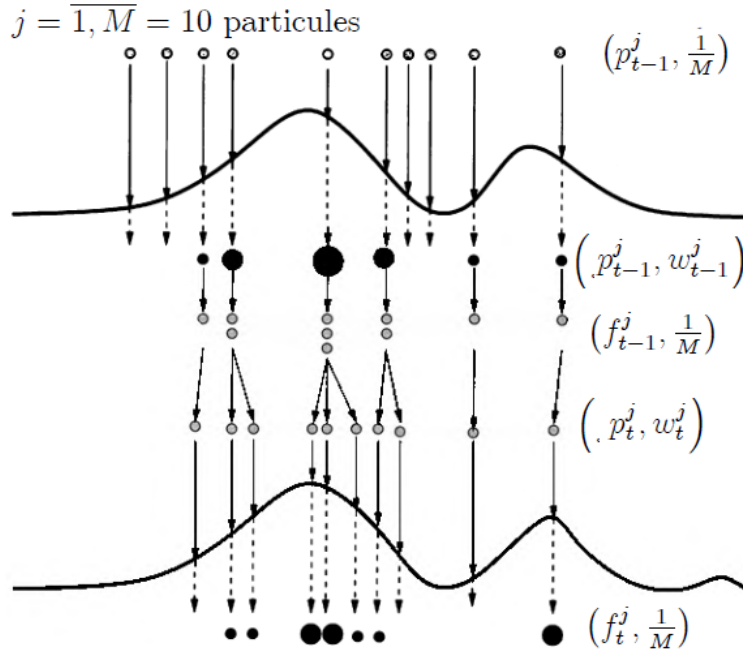


Figure 1.4. Principe du bootstrap²

Table 1.1 illustration du *SIS* et *SIR*

l'instant	la densité conditionnelle	l'approximation	remarque	
$t - 1$	$p(x_{t-1} Y_{t-2})$	$\left\{ \left(p_{t-1}^j, \frac{1}{M} \right), j = \overline{1, M} \right\}$	filtrage sans rééchantillonnage	filtrage avec rééchantillonnage
$t - 1$	$p(x_{t-1} Y_{t-1})$	$\left\{ \left(\tilde{f}_{t-1}^j = p_{t-1}^j, w_{t-1}^j \right), j = \overline{1, M} \right\}$		
$t - 1$	$p(x_{t-1} Y_{t-1})$	$\left\{ \left(f_{t-1}^j, \frac{1}{M} \right), j = \overline{1, M} \right\}$	étape de rééchantillonnage	équivalent à <i>SIR</i>
t	$p(x_t Y_{t-1})$	$\left\{ \left(p_t^j, \frac{1}{M} \right), j = \overline{1, M} \right\}$	cette étape n'est pas illustrée sur la Figure 1.4	
t	$p(x_t Y_t)$	$\left\{ \left(\tilde{f}_t^j = p_t^j, w_t^j \right), j = \overline{1, M} \right\}$		
t	$p(x_t Y_t)$	$\left\{ \left(f_t^j, \frac{1}{M} \right), j = \overline{1, M} \right\}$		

Notons que le symbole « $\tilde{}$ » surmontant une particule indique que la particule en question n'a pas encore subi l'étape de rééchantillonnage. Cette notation a pour vocation de faciliter la compréhension de la chronologie des différents processus qui composent une récurrence d'un filtre particulaire. A l'instant $t - 1$, les particules, dont les poids d'importance sont semblables $\left\{ p_t^{(j)}, 1/M \right\}_{j=\overline{1, M}}$, sont distribuées de telle manière à fournir une approximation de la distribution a posteriori prédite $p(x_{t-1}|Y_{t-2})$.

A l'aide de la mesure à $t - 1$, on procède à la mise à jour du poids d'importance pour chacune des particules. De cette manière, on obtient un essaim de particules pondérées $\left\{ \tilde{f}_{t-1}^{(j)} = p_{t-1}^{(j)}, w_{t-1}^{(j)} \right\}_{j=\overline{1, M}}$ qui fournit une estimation empirique de la distribution $p(x_{t-1}|Y_{t-1})$. Puis, l'étape de rééchantillonnage duplique les particules de forte vraisemblance, pour obtenir une approximation de la distribution $p(x_{t-1}|Y_{t-1})$ à partir des particules non pondérées $\left\{ f_{t-1}^{(j)}, 1/M \right\}_{j=\overline{1, M}}$ (voir la Figure 1.4).

²Figure extraite de Doucet et al. (2001, p. 12)

1.4.5 Lissage particulière

L'objectif de l'étape de lissage est de fournir des particules lissées $\left\{s_t^{(j)}, w_t^{(j)}\right\}_{j=\overline{1, M}}$ pour approximer $p(x_t|Y_n)$. Théoriquement, la trajectoire des états $\{s_1, \dots, s_n\}$ à partir de $p(x_1, \dots, x_n|Y_n)$ peut être obtenue en utilisant les particules filtrées. Godsil et al. (2004) ont suggéré une méthode de lissage particulière, basée sur la simulation en arrière (particle smoother using backwards simulation). Contrairement aux autres méthodes de lissage, cette dernière ne souffre pas du problème de dégénérescence et elle concerne l'ensemble des états de la trajectoire $\{x_0, \dots, x_n\}$ de la densité conjointe $p(x_0, \dots, x_n|Y_n)$ et pas seulement la densité marginale, $p(x_t|Y_n)$. Le lissage particulière en utilisant la simulation en arrière suppose que le filtrage a déjà été effectué de sorte que les particules et les poids associés $\left\{f_t^{(j)}, w_t^{(j)}\right\}_{j=\overline{1, M}}$ approximent bien la densité $p(x_t|Y_t)$ par $\sum_{j=1}^M w_t^{(j)} \delta_{f_t^{(j)}}(x_t)$. Cette méthode peut être justifiée comme suit,

$$p(x_1, \dots, x_n|Y_n) = p(x_n|Y_n) \prod_{t=1}^n p(x_t|x_{t+1}, \dots, x_n, Y_n),$$

et

$$p(x_t|x_{t+1}, \dots, x_n, Y_n) = p(x_t|x_{t+1}, Y_t) = \frac{p(x_t|Y_t) p(x_{t+1}|x_t)}{p(x_{t+1}|Y_t)} \propto p(x_t|Y_t) p(x_{t+1}|x_t).$$

Dans cette méthode, les particules approximant $p(x_t|Y_t)$, sont supposées disponibles (obtenues par l'étape de filtrage), donc nous obtenons immédiatement l'approximation des particules modifiées suivante

$$p(x_t|s_{t+1}, \dots, s_n, Y_n) \approx \sum_{i=1}^M w_{t|t+1}^{(i)} \delta_{f_t^{(i)}}(x_t),$$

avec les poids modifiés

$$w_{t|t+1}^{(i)} = \frac{w_t^{(i)} p(x_{t+1}|f_t^{(j)})}{\sum_{k=1}^M w_t^{(k)} p(x_{t+1}|f_t^{(k)})}.$$

Alors, l'idée est la suivante : étant donné $\left\{s_{t+1}^{(j)}, \dots, s_n^{(j)}\right\}$ une trajectoire de $p(x_{t+1}, \dots, x_n|Y_n)$, il est possible d'avoir $s_t^{(j)}$ générée à partir de $p(x_t|s_{t+1}, \dots, s_n, Y_n) = p(x_t|s_{t+1}, Y_n)$ et on obtient donc $\left\{s_t^{(j)}, s_{t+1}^{(j)}, \dots, s_n^{(j)}\right\}$ qui est une réalisation de $p(x_t, \dots, x_n|Y_n)$. Donc, on va commencer par $t = n$, où on génère $s_n^{(j)}$ à partir de $p(x_n|Y_n)$ qui est déjà approximée par $\left\{f_t^{(i)}, w_t^{(i)}\right\}_{i=\overline{1, M}}$. Autrement dit, on doit faire un rééchantillonnage de $\left\{f_t^{(i)}\right\}_{i=\overline{1, M}}$ avec les poids $\left\{w_t^{(i)}\right\}_{i=\overline{1, M}}$. Ensuite, en répétant cette procédure séquentiellement, où on va générer pour chaque instant t , s_t à partir $p(x_t|s_{t+1}, Y_n)$ qui doit être approximée par $\left\{f_t^{(i)}, w_{t|t+1}^{(i)}\right\}_{i=\overline{1, M}}$, d'où la nécessité de calculer les poids lissés $\left\{w_{t|t+1}^{(i)}\right\}_{i=\overline{1, M}}$. On obtient à la fin les particules lissées $\left\{s_t^{(j)}\right\}_{t=1, \dots, n}$ qui est une réalisation aléatoire de $p(X_n|Y_n)$. L'algorithme suivant présente les différentes phases du fonctionnement du lissage particulière.

Algorithme 1.4 (Lissage particulière (Godsill et al., 2004))

1. Pour $j = 1, \dots, M$,

(a) Pour $t = n$, choisir $s_n^{(j)} = f_n^{(i)}$ avec une probabilité $w_n^{(i)}$, $i = \overline{1, M}$.

(b) Pour $t = n - 1, \dots, 1$,

i. Pour $i = \overline{1, M}$, calculer $w_{t|t+1}^{(i)} \propto w_t^{(i)} p(s_{t+1}^{(j)} | f_t^{(i)})$.

ii. Choisir $s_t^{(j)} = f_t^{(i)}$ avec une probabilité $w_{t|t+1}^{(i)}$.

2. Enfin, la distribution empirique de $\{s_1^{(j)}, s_2^{(j)}, \dots, s_n^{(j)}\}_{j=\overline{1, M}}$ est une approximation de $p(X_n | Y_n)$.

La complexité de calcul pour chaque réalisation est $O(Mn)$, contrairement au $O(M^2n)$ requis pour les procédures de lissage marginal (Kitagawa, 1996 ; Doucet et al., 2000 ; Hürzeler et Künsch 1998).

1.4.6 Filtre particulière Rao-Blackwellisé

Le fondement de la technique dite de Rao-Blackwellisation (Casella et Robert, 1996) réside dans l'exploitation de la linéarité de certaines composantes du vecteur d'état du modèle, de façon à réduire la variance de l'erreur. L'estimation du vecteur d'état, dans cette méthode, est effectuée en hybridant un filtre particulière avec une bande de filtres de Kalman ; seule la partie non linéaire est traitée avec le filtrage particulière, le reste du calcul est réalisé analytiquement par le filtre de Kalman. Un tel filtre améliore l'efficacité de l'inférence, car une partie du problème est résolu de façon optimale (voir Doucet et al., 2000).

1.4.7 Synthèse

L'objet de ce chapitre était une brève présentation de deux méthodes d'estimation des variables d'états d'un modèle espace d'états. Toutefois, nous tenons à préciser les points, indiquant les choix retenus dans cette étude, et qui sont

- l'utilisation de la première définition pour la mesure de rééchantillonnage,
- l'emploi du rééchantillonnage multinomial,
- le choix de la loi d'évolution de l'état comme densité d'importance.

Ainsi et pour plus de détails, en particulier les démonstrations de théorèmes ou encore les études de convergence, nous invitons le lecteur à se référer à des études plus poussées (Gordon et al., 1993 ; doucet et al., 2001 ; Godsill et al., 2004 ; Doucet et al 2012).

Chapitre 2

Modèles à Volatilité Stochastique Univariés à Coefficients Périodiques

2.1 Introduction

La littérature sur la modélisation et la prévision des phénomènes stochastiques ayant une structure d'autocorrélation périodique est volumineuse. En effet, des modèles de séries chronologiques périodiques ont été largement utilisés au cours des dernières décennies pour décrire de nombreuses séries temporelles présentant des dynamiques périodiques, rencontrées dans une grande variété de domaines scientifiques. On peut citer à titre d'exemple l'environnement (Vecchia, 1985 ; Jiménez et al., 1989), l'économie (Franses, 1996 ; Franses et Paap, 2004), la finance (Bollerslev et Ghysels, 1996 ; Franses et Paap, 2000 ; Bentarzi et Hamdi, 2008 ; Hamdi et Souam, 2017). C'est ainsi que de nombreux modèles linéaires périodiques ont été introduits ; parmi lesquels, nous trouvons les modèles Périodiques Auto-Régressive Moyenne Mobile (*PARMA*). Ces derniers ont été extensivement étudiés par de nombreux auteurs (voir par exemple Vecchia, 1985 ; Jiménez et al., 1989 ; Bentarzi et Hallin, 1994 ; Boshnakov, 1996 ; Ula et Smadi, 1997 ; Lund et Basawa, 2000, Basawa et Lund, 2001 ; Shao et Lund, 2004 ; Bentarzi et Aknouche, 2005 ; Bentarzi et al., 2008 ; Aknouche et al., 2008 ; Aknouche et Hamdi, 2009 ; Hamdi, 2012 ; Guerbyenne et Hamdi, 2015 et bien d'autres).

De son côté, l'approche d'hétéroscédasticité conditionnelle offre des outils d'analyse et de conception puissants pour la modélisation périodique des séries chronologiques stationnaires. On peut citer, en particulier, la classe des modèles AutoRégressifs Conditionnellement Hétéroscédastiques généralisés à coefficients périodiques dans le temps (*PGARCH*) introduite par Bollerslev et Ghysels (1996). Celle-ci constitue, sans doute, la classe la plus importante pour la modélisation des séries chronologiques financières périodiques. De plus, elle s'est révélée appropriée pour capter la périodicité dans la variance conditionnelle, propriété qui ne peut pas être expliquée par les formulations *GARCH* classiques.

Une autre catégorie de modèles statistiques, à savoir la classe des modèles à volatilité stochastique (*SV*) introduite par Taylor (1982), est considérée comme une alternative satisfaisante aux modèles de type-*ARCH*. Cette approche est plus flexible et capte mieux les variations dans la volatilité des actifs

en permettant à la variance des rendements d'être un processus aléatoire non observé, contrairement aux modèles *ARCH* et leurs généralisations *GARCH*, où la volatilité est une fonction des observations retardées et/ou des volatilités passées. En d'autres termes, ces modèles supposent que la variance future est connue de façon certaine ce qui n'est pas tout-à-fait cohérent avec la réalité. La modélisation *SV* est motivée par la conviction que les changements dans la variance sont causés par des conditions économiques et politiques, et la volatilité ne peut pas être prédite exactement uniquement sur les rendements passés.

La version la plus populaire du modèle *SV* considère que le logarithme de la volatilité suit un processus autorégressif de premier ordre. Notons que la modélisation *SV* permet à la log-volatilité d'évoluer au cours du temps, ce qui assure la positivité de la variance du processus sans avoir besoin de contraintes supplémentaires, comme c'est le cas dans les modèles *ARCH/GARCH*.

Par ailleurs, de nombreuses extensions du modèle *SV* classique, ont été proposées dans la littérature économétrique pour capturer différentes caractéristiques de séries chronologiques. Cependant, la plupart des formulations proposées traitent les paramètres du modèle *SV* comme des constantes en fonction du temps. Elles ne peuvent, de ce fait, expliquer, de manière adéquate, les séries chronologiques dont la volatilité affiche une certaine périodicité et qui ne peuvent pas être expliquées par les modèles *SV* avec des paramètres invariants dans le temps. Il est donc possible de considérer, similairement aux modèles *GARCH* périodiques (Bollerslev et Ghysels, 1996), une formulation à coefficients périodiques dans le temps pour l'équation de la log-volatilité.

Le but de ce chapitre est d'étudier une nouvelle classe de modèles à volatilité stochastique à coefficients périodiques, introduite dans un document non publié par Aknouche et *al.* (2007), pour modéliser, à la fois, la variabilité instantanée et la périodicité à travers une meilleure formulation. Ce modèle a été appelé Volatilité Stochastique AutoRégressif Périodique (*PAR-SV*) plutôt que Volatilité Stochastique Périodique (*PSV*), puisque la log-volatilité est définie comme un processus *AR* périodique d'ordre un. Il diffère du modèle *PSV* proposé par Tsiakas (2006). Notons que dans la formulation de Tsiakas, les paramètres sont exprimés comme une combinaison de sinus et de cosinus avec des variables indicatrices appropriées. Cette représentation ne décrit, toutefois, qu'un cas particulier de saisonnalité déterministe (Ghysels et Osborn, 2001) et ne tient pas compte de la volatilité périodiquement corrélée. Notons également que la classe des modèles *PAR-SV* a été étudiée récemment et indépendamment de notre travail, par Aknouche (2017).

Ce chapitre est organisé de la manière suivante. La Section 2.2 propose quelques définitions préliminaires sur les modèles *PAR-SV* et leur représentation espace d'états linéaire périodique. Dans la Section 2.3, nous proposons certaines propriétés probabilistes du modèle étudié, à savoir la stationnarité périodique stricte et de second ordre. Dans la Section 2.4, nous décrivons deux méthodes simples pour l'estimation du modèle. La mise en œuvre est basée sur le filtre de Kalman périodique et sur l'algorithme d'Expectation-Maximisation (*EM*) combiné avec le filtrage et le lissage particuliers. Finalement, dans la Section 2.5, nous proposons une étude de simulation pour examiner la performance des méthodes d'estimation proposées et une application du modèle *PAR-SV* à la série journalière Euro/Dinar Algérien (*EUR/DZD*) et à la série journalière Dollar Américain par rapport au Dinar Algérien (*USD/DZD*)

2.2 Modèle à Volatilité Stochastique AutoRégressif Périodique

Il existe dans la littérature économétrique plusieurs variantes du modèle à volatilité stochastique univarié, dont la première a été introduite par Taylor en 1982. Cependant, nous présentons le modèle utilisé par Platanioti et al. (2005). Considérons le processus $\{\varepsilon_t; t \in \mathbb{Z}\}$ satisfaisant la représentation espace d'états suivante

$$\begin{cases} \varepsilon_t = \eta_t \exp \left\{ \frac{1}{2} x_t \right\}, \\ x_t = \alpha + \beta x_{t-1} + Q e_t, \end{cases} \quad t \in \mathbb{Z}, \quad (2.1)$$

où ε_t est le rendement d'un actif au temps t et x_t est la log-volatilité correspondante. Les coefficients α, β et Q sont respectivement un facteur d'échelle pour la volatilité, le coefficient de persistance et une mesure d'amplitude des oscillations de x_t autour de sa moyenne. Enfin, (η_t) et (e_t) sont deux séquences de variables aléatoires *i.i.d.* centrées réduites mutuellement indépendantes. Dans ce modèle, $\{\varepsilon_t; t \in \mathbb{Z}\}$ décrit la série temporelle sous-jacente et la variable non observée x_t est connue comme la log-volatilité du processus $\{\varepsilon_t; t \in \mathbb{Z}\}$.

Pour modéliser à la fois la variabilité instantanée et la périodicité à travers une meilleure formulation, Aknouche et al. (2007) ont proposé d'ajouter plus de flexibilité à la modélisation *SV* standard, en introduisant la classe des modèles à volatilité stochastique autorégressifs périodiques. Rappelons qu'un processus stochastique $\{\varepsilon_t; t \in \mathbb{Z}\}$ admet une représentation *Periodic AutoRegressive Stochastic Volatility (PAR-SV)* de période S , s'il est solution de l'équation aux différences stochastique suivante

$$\begin{cases} \varepsilon_t = \eta_t \exp \left\{ \frac{1}{2} x_t \right\}, \\ x_t = \alpha_t + \beta_t x_{t-1} + Q_t e_t, \end{cases} \quad t \in \mathbb{Z}, \quad (2.2)$$

où (η_t) et (e_t) sont deux suites mutuellement indépendantes de variables aléatoires *i.i.d.* centrées réduites, et les paramètres α_t, β_t et Q_t sont périodiques en t de période S (i.e., $\alpha_{t+\tau S} = \alpha_t, \beta_{t+\tau S} = \beta_t$ et $Q_{t+\tau S} = Q_t$). Notons que le modèle *PAR-SV* défini par (2.2) peut être généralisé de sorte que x_t soit un processus *PARMA*. Notons également que dans le modèle (2.2), la volatilité devient, par le biais de l'introduction d'un autre processus d'innovation injecté dans la structure de la variance conditionnelle, une variable non observable. Par conséquent, la classe des modèles *PAR-SV* est plus flexible que celle des modèles *PGARCH* (voir par exemple Kim et al., 1998, pour le cas non périodique). En effet, lorsque la partie stochastique de la volatilité, e_t , n'apparaît pas, le modèle (2.2) ne se réduit pas à un modèle *PARCH(1)*, mais à un modèle *PGARCH(0, 1)*. Donc, la différence dans la modélisation de la volatilité est importante entre les deux approches *PGARCH* et *PAR-SV*.

Réécrivons le modèle *PAR-SV* défini dans (2.2), en faisant ressortir la périodicité comme suit

$$\begin{cases} \varepsilon_{s+\tau S} = \eta_{s+\tau S} \exp \left\{ \frac{1}{2} x_{s+\tau S} \right\}, \\ x_{s+\tau S} = \alpha_s + \beta_s x_{s+\tau S-1} + Q_s e_{s+\tau S}, \end{cases} \quad \tau \in \mathbb{Z}, \quad 1 \leq s \leq S. \quad (2.3)$$

La linéarisation de l'équation d'observation par passage au logarithme de $\varepsilon_{s+\tau S}^2$, conduit à la représentation espace d'états suivante

$$\begin{cases} Y_{s+\tau S} = x_{s+\tau S} + d + u_{s+\tau S}, \\ x_{s+\tau S} = \alpha_s + \beta_s x_{s+\tau S-1} + Q_s e_{s+\tau S}, \end{cases} \quad \tau \in \mathbb{Z}, \quad 1 \leq s \leq S, \quad (2.4)$$

où $Y_{s+\tau S} = \log(\varepsilon_{s+\tau S}^2)$, $d = \mathbb{E}[\log(\eta_{s+\tau S}^2)]$ et $u_{s+\tau S} = \log(\eta_{s+\tau S}^2) - \mathbb{E}[\log(\eta_{s+\tau S}^2)]$. Si $\eta_{s+\tau S}$ est une variable gaussienne, comme il est souvent supposé dans la littérature, la variable $u_{s+\tau S}$ est distribuée selon une densité log de *Chi-deux* ($\log \chi^2$) avec un degré de liberté (voir par exemple Harvey et al., 1994). Cette densité est donnée par

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\exp(x) - x)\right), \quad x \in \mathbb{R}.$$

La densité de $\log(\eta_{s+\tau S}^2)$ est fortement asymétrique et présente une longue queue à gauche. La moyenne et la variance de $\log(\eta_{s+\tau S}^2)$ peuvent être approximées respectivement par -1.2749 et $\pi^2/2$. Une comparaison entre la densité de $\log(\eta_{s+\tau S}^2)$ et la loi normale $\mathcal{N}(-1.2749, \pi^2/2)$ est présentée dans la Figure 2.1. De toute évidence, cette approche qui repose sur la linéarisation de l'équation d'observation est une approximation du modèle *PAR-SV*.

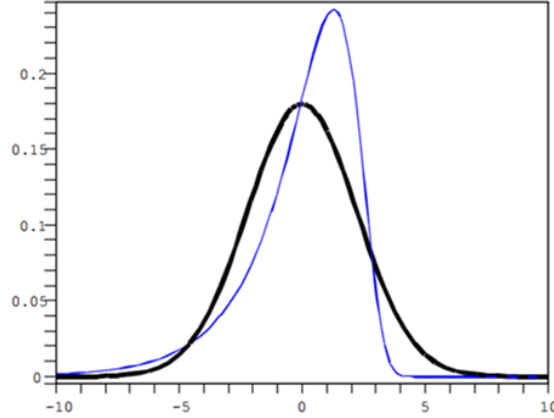


Figure 2.1. Comparaison de la densité $\log \chi_1^2$ (trait fin) avec la densité normale $\mathcal{N}(0, \pi^2/2)$ (trait épais)

2.3 Étude Probabiliste du modèle *PAR-SV*

2.3.1 Stationnarité périodique

Rappelons tout d'abord les deux notions standards de la stationnarité périodique. Le processus $\{y_t, t \in \mathbb{Z}\}$ est dit périodiquement fortement stationnaire (ou périodiquement stationnaire au sens strict) s'il existe un entier positif S tel que pour tout entier positif k et pour toute suite t_1, t_2, \dots, t_k , la distribution conjointe de $(y_{t_1}, y_{t_2}, \dots, y_{t_k})$ est la même que celle de $(y_{t_1+S}, y_{t_2+S}, \dots, y_{t_k+S})$. Le plus petit S vérifiant cette propriété est appelé période du processus.

Le processus du second ordre $\{y_t, t \in \mathbb{Z}\}$ est dit périodiquement faiblement stationnaire (ou périodiquement stationnaire au second ordre) de période S , si sa moyenne et sa fonction d'autocovariance sont périodiques dans le temps, i.e. $\mu_{t+\tau S} = \mu_t$ et $\gamma_h^{(t+\tau S)} = \gamma_h^{(t)}$, $\forall t, h, \tau \in \mathbb{Z}$.

Pour une analyse statistique il est important de chercher des conditions assurant la stationnarité périodique (au sens strict et au second ordre). Considérons un processus $\{\varepsilon_t, t \in \mathbb{Z}\}$ qui suit un modèle *PAR-SV*. On s'intéresse particulièrement aux solutions non anticipatives du modèle (2.4), c'est-à-dire aux processus ε_t , tels que ε_t est une fonction mesurable des variables η_{t-s} , $s \geq 0$.

Théorème 2.1 (Boussaha et Hamdi, 2017) *Une condition nécessaire et suffisante d'existence d'une solution périodiquement strictement stationnaire (p.s.s.) de l'équation aux différences stochastique (2.2), est donnée par $\left| \prod_{v=1}^S \beta_v \right| < 1$. De plus, dans le cas où e_t est normalement distribuée, les conditions de stationnarité périodique au second ordre et stricte coïncident.*

Preuve. Il est clair, au vu de la forme multiplicative de la première équation de (2.2) que les conditions d'existence d'une solution p.s.s. sont équivalentes aux conditions d'existence d'une solution causale p.s.s. d'un modèle *PAR*(1) donnée par

$$x_{s+\tau S} = \alpha_s + \beta_s x_{s+\tau S-1} + Q_s e_{s+\tau S},$$

qui se résumant par $\left| \prod_{v=1}^S \beta_v \right| < 1$. En effet, sous cette condition, la log-volatilité x_t admet la représentation causale suivante

$$x_{s+\tau S} = \mu_x^{(s)} + \sum_{l \geq 0} \psi_{s,l} Q_{s-l} e_{s+\tau S-l},$$

dont les coefficients S -périodiques $\psi_{s,l}$ sont absolument sommables, i.e. $\sum_{l \geq 0} |\psi_{s,l}| < \infty$, pour chaque période $s = \overline{1, S}$. Ces coefficients peuvent être obtenus récursivement en fonction des paramètres du modèle, via la relation suivante (voir Lund and Basawa, 2000)

$$\begin{cases} \psi_{s,0} = 1, \\ \psi_{s,l} = \beta_s \psi_{s-1,l-1} = \prod_{k=0}^{l-1} \beta_{s-k}, \end{cases}, \quad s = \overline{1, S}.$$

Il est clair que la moyenne périodique inconditionnelle $\mu_x^{(s)} := \mathbb{E}(x_{s+\tau S})$ donnée par

$$\mu_x^{(s)} = \frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{s-j} \right) \alpha_{s-i}}{1 - \prod_{v=1}^S \beta_v}, \quad \text{pour } s = \overline{1, S}. \quad (2.5)$$

Ainsi, nous avons

$$\mathbb{E}(\varepsilon_{s+\tau S}^2) = \exp \{ \mu_x^{(s)} \} \mathbb{E} \left(\prod_{l \geq 0} \exp \{ \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \} \right).$$

De plus, il est possible de prouver que

$$\begin{aligned} \mathbb{E} \left(\prod_{l \geq 0} \exp \{ \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \} \right) &= \prod_{l \geq 0} \mathbb{E} \left(\exp \{ \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \} \right) \\ &= \prod_{l \geq 0} M_e(\psi_{s,l} Q_{s-l}) \\ &= \exp \left(\frac{1}{2} \sum_{l \geq 0} \left(\prod_{k=0}^{l-1} \beta_{s-k}^2 \right) Q_{s-l}^2 \right) \\ &\leq \exp \left(\frac{1}{2} Q^2 \sum_{l \geq 0} \left(\prod_{k=0}^{l-1} \beta_{s-k}^2 \right) \right) < \infty, \end{aligned}$$

où $Q = \max_{s=\overline{1, S}} Q_s$ et $M_e(z) = \mathbb{E}(\exp \{ z e_t \})$ désigne la fonction génératrice des moments de e_t .

Soit maintenant

$$U_{s+\tau S, L} = \prod_{l=0}^L Z_{s+\tau S, l} = \exp \left\{ \sum_{l=0}^L \frac{1}{2} \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \right\},$$

où $Z_{s+\tau S,l} = \exp \left\{ \frac{1}{2} \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \right\}$, $s \in \{1, \dots, S\}$, $\tau \in \mathbb{Z}$ et $L \in \mathbb{N}$. Notons par $\mathcal{F}_{s+\tau S,L}$ la σ -algèbre engendrée par $Z_{s+\tau S,0}, Z_{s+\tau S,1}, \dots, Z_{s+\tau S,L}$, nous pouvons voir, à partir de l'inégalité de Jensen, que pour tout $L \in \mathbb{N}$

$$\begin{aligned} \mathbb{E}(U_{s+\tau S,L+1} | \mathcal{F}_{s+\tau S,L}) &= U_{s+\tau S,L} \mathbb{E} \left(\exp \left\{ \frac{1}{2} \psi_{s,L+1} Q_{s-L-1} e_{s+\tau S-L-1} \right\} \right) \\ &\geq U_{s+\tau S,L} \exp \left\{ \mathbb{E} \left(\frac{1}{2} \psi_{s,L+1} Q_{s-L-1} e_{s+\tau S-L-1} \right) \right\} = U_{s+\tau S,L}, \end{aligned}$$

et donc $(U_{s+\tau S,L})_{L \geq 0}$ est une sous-martingale positive pour tout $s = \overline{1, S}$ et $\tau \in \mathbb{Z}$. De plus, elle est carrée intégrable, puisque

$$\begin{aligned} \mathbb{E}(U_{s+\tau S,L}^2) &= \mathbb{E} \left(\prod_{l=0}^L \exp \{ \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \} \right) = \exp \left\{ \frac{1}{2} \sum_{l=0}^L \psi_{s,l}^2 Q_{s-l}^2 \right\} \\ &\leq \exp \left\{ \frac{1}{2} \sum_{l \geq 0} \psi_{s,l}^2 Q_{s-l}^2 \right\} < \infty. \end{aligned}$$

Posons $S_{s+\tau S,L} = \sup_{k \leq L} U_{s+\tau S,k}$ et $S_{s+\tau S,\infty} = \sup_k U_{s+\tau S,k}$, nous avons à partir de l'inégalité de Doob

$$\mathbb{E}(S_{s+\tau S,L}) \leq [\mathbb{E}(S_{s+\tau S,L}^2)]^{1/2} \leq 2 [\mathbb{E}(U_{s+\tau S,L}^2)]^{1/2} \leq 2 \exp \left\{ \frac{1}{4} \sum_{l \geq 0} \psi_{s,l}^2 Q_{s-l}^2 \right\} < \infty,$$

et d'après le Lemme de Fatou, nous avons

$$\mathbb{E}(S_{s+\tau S,\infty}) \leq \exp \left\{ \frac{1}{2} \sum_{l \geq 0} \psi_{s,l}^2 Q_{s-l}^2 \right\} < \infty,$$

ainsi, nous pouvons utiliser le théorème de convergence dominée, nous obtenons

$$\mathbb{E} \left(\lim_{L \rightarrow \infty} U_{s+\tau S,L} \right) = \lim_{L \rightarrow \infty} \mathbb{E}(U_{s+\tau S,L}).$$

En conséquence, si (e_t) est normalement distribuée, alors, la condition $\left| \prod_{v=1}^S \beta_v \right| < 1$, est nécessaire et suffisante pour assurer l'existence d'une solution *p.s.s.* de (2.2) admettant un moment d'ordre 2 fini. ■

Remarque 2.1 *Dans le cas général et contrairement à ce qui est donné dans le théorème précédent (sous l'hypothèse de la normalité faite sur la distribution de (e_t) , la condition $\left| \prod_{v=1}^S \beta_v \right| < 1$, n'est pas suffisante pour assurer la stationnarité périodique au second ordre de (2.2). Ce problème a été récemment traité dans Aknouche (2017) (voir Théorème 2.3). Notons que les hypothèses (2.6a), (2.6b) et (2.6c) supposées par Aknouche (2017) pour établir l'existence du moment d'ordre 2 peuvent être remplacées par l'hypothèse suivante*

$$\prod_{l \geq 0} M_e(\psi_{s,l} Q_{s-l}) < \infty, \text{ pour } s = \overline{1, S}.$$

En effet, pour montrer que $\mathbb{E}(\varepsilon_{s+\tau S}^2)$ est fini, il suffit de prouver que $\mathbb{E} \left(\prod_{l \geq 0} \exp \{ \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \} \right) < \infty$, ce qui est vrai selon le Lemme de Fatou car $(U_{s+\tau S,L})_{L \geq 0}$ est une suite de variables aléatoires intégrables positives qui convergent presque sûrement.

Remarque 2.2 Dans le cas général, on peut même montrer que si $\prod_{l \geq 0} \mathbb{E} \left(\exp \left\{ \frac{r}{2} \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \right\} \right) < \infty$, où r est un entier positif, alors nous avons

$$\mathbb{E} \left(\prod_{l \geq 0} \exp \left\{ \frac{r}{2} \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \right\} \right) = \prod_{l \geq 0} \mathbb{E} \left(\exp \left\{ \frac{r}{2} \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \right\} \right).$$

Il suffit de procéder exactement de la même manière que dans la preuve du cas gaussien avec $r = 2$, mais cette fois la sous-martingale $(U_{s+\tau S, L})_{L \geq 0}$ est définie par

$$U_{s+\tau S, L} = \prod_{l=0}^L Z_{s+\tau S, l},$$

où $Z_{s+\tau S, l} = \exp \left\{ \frac{r}{4} \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \right\}$, pour $s \in \{1, \dots, S\}$, $\tau \in \mathbb{Z}$ et $L \in \mathbb{N}$.

2.3.2 Moments et structure dynamique d'un modèle PAR-SV

Notons que si $\left| \prod_{s=1}^S \beta_s \right| < 1$, le processus $\{x_t; t \in \mathbb{Z}\}$ est *p.s.s.* de moyenne périodique inconditionnelle $\mu_x^{(t)}$ et de variance périodique inconditionnelle

$$\sigma_x^{2(s)} := \text{var} (x_{s+\tau S}) = \frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{s-j}^2 \right) \left[\left(\alpha_{s-i} + \beta_{s-i} \mu_x^{(s-i-1)} - \mu_x^{(s-i)} \right)^2 + Q_{s-i}^2 \right]}{1 - \prod_{v=1}^S \beta_v^2}, \quad (2.6)$$

pour $s = \overline{1, S}$.

Moments d'ordre r d'un modèle PAR-SV

Nous pouvons affirmer, à l'aide du Théorème 2.1, que lorsque $\left| \prod_{s=1}^S \beta_s \right| < 1$, le processus $\{\varepsilon_t; t \in \mathbb{Z}\}$ est *p.s.s.* et par conséquent, les moments de ce processus, s'ils existent, sont périodiques dans le temps. Notons que si nous faisons une hypothèse de symétrie sur la loi de η_t , les moments d'ordre impair de (η_t) , lorsqu'ils existent, sont nuls. De plus, si $\mathbb{E} (|\eta_t|^r) < \infty$, où r est un entier positif, les moments d'ordre pair de (ε_t) , peuvent être calculés en utilisant les résultats standards de la distribution log-normale de la variable $\exp \{x_t\}$. En effet, nous avons le résultat suivant.

Théorème 2.2 (Boussaha et Hamdi, 2017) Soit $\{\varepsilon_t; t \in \mathbb{Z}\}$ une solution *p.s.s.* de (2.2), avec $\mathbb{E} (|\eta_t|^r) < \infty$, où r est un entier positif. Alors une condition suffisante pour que $\mathbb{E} (\varepsilon_t^r)$ soit finie est donnée par

$$\prod_{l \geq 0} M_e \left(\frac{r}{2} \psi_{s,l} Q_{s-l} \right) < \infty, \quad \text{pour } s = \overline{1, S},$$

où $M_e(z) = \mathbb{E} [\exp(z e_t)]$ désigne la fonction génératrice des moments de e_t . De plus, l'expression explicite des moments d'ordre r de ε_t , est donnée par

$$\mu_{\varepsilon^r}^{(s)} := \mathbb{E} (\varepsilon_{s+\tau S}^r) = \mathbb{E} (\eta_{s+\tau S}^r) \exp \left\{ \frac{r \sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{s-j} \right) \alpha_{s-i}}{2 \left(1 - \prod_{v=1}^S \beta_v \right)} \right\} \prod_{l \geq 0} M_e \left(\frac{r}{2} \psi_{s,l} Q_{s-l} \right).$$

Preuve. D'après l'hypothèse d'indépendance sur les suites (η_t) et (e_t) , nous avons pour chaque $\tau \in \mathbb{Z}$ et $1 \leq s \leq S$,

$$\begin{aligned} \mathbb{E}(\varepsilon_{s+\tau S}^r) &= \mathbb{E}(\eta_{s+\tau S}^r) \exp\left\{\frac{r}{2}\mu_x^{(s)}\right\} \mathbb{E}\left(\exp\left\{\frac{r}{2}\sum_{l \geq 0} \psi_{s,l} Q_{s-l} e_{s+\tau S-l}\right\}\right) \\ &= \mathbb{E}(\eta_{s+\tau S}^r) \exp\left\{\frac{r}{2}\mu_x^{(s)}\right\} \prod_{l \geq 0} M_e\left(\frac{r}{2}\psi_{s,l} Q_{s-l}\right) \\ &= \mathbb{E}(\eta_{s+\tau S}^r) \exp\left\{\frac{r \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}) \alpha_{s-i}}{2(1-\prod_{v=1}^S \beta_v)}\right\} \prod_{l \geq 0} M_e\left(\frac{r}{2}\psi_{s,l} Q_{s-l}\right). \end{aligned}$$

Il en découle facilement qu'une condition suffisante pour que $\mathbb{E}(\varepsilon_{s+\tau S}^r)$ soit finie est

$$\prod_{l \geq 0} M_e\left(\frac{r}{2}\psi_{s,l} Q_{s-l}\right) < \infty.$$

■

Remarque 2.3 Si de plus $\mathbb{E}(\eta_t^r) \neq 0$, la condition $\prod_{l \geq 0} M_e\left(\frac{r}{2}\psi_{s,l} Q_{s-l}\right) < \infty$, pour $s = \overline{1, S}$, devient une condition nécessaire et suffisante pour que $\mathbb{E}(\varepsilon_t^r)$ soit finie.

Remarque 2.4 Sous l'hypothèse que $\mathbb{E}(|\eta_t|^r) < \infty$, et aussi la normalité de (e_t) , la condition du Théoreme 2.2 est vérifiée. En effet, nous avons pour $1 \leq s \leq S$,

$$\begin{aligned} \prod_{l \geq 0} M_e\left(\frac{r}{2}\psi_{s,l} Q_{s-l}\right) &= \prod_{l \geq 0} \exp\left\{\frac{r^2 \psi_{s,l}^2 Q_{s-l}^2}{8}\right\} = \exp\left\{\frac{r^2 \sum_{l \geq 0} \psi_{s,l}^2 Q_{s-l}^2}{8}\right\} \\ &= \exp\left\{\frac{r^2 \sum_{i=0}^{S-1} \sum_{\tau=0}^{\infty} (\prod_{v=1}^S \beta_v)^{2\tau} (\prod_{j=0}^{i-1} \beta_{s-j}^2) Q_{s-i}^2}{8}\right\} \\ &= \exp\left\{\frac{r^2 \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}^2) Q_{s-i}^2}{8(1-\prod_{v=1}^S \beta_v^2)}\right\} < \infty. \end{aligned}$$

Ainsi, l'expression explicite des moments d'ordre r de $\varepsilon_{s+\tau S}$ est donnée par

$$\begin{aligned} \mathbb{E}(\varepsilon_{s+\tau S}^r) &= \mathbb{E}(\eta_{s+\tau S}^r) \exp\left\{\frac{r \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}) \alpha_{s-i}}{2(1-\prod_{v=1}^S \beta_v)}\right\} \prod_{l \geq 0} M_e\left(\frac{r}{2}\psi_{s,l} Q_{s-l}\right) \\ &= \mathbb{E}(\eta_{s+\tau S}^r) \exp\left\{\frac{r(1+\prod_{v=1}^S \beta_v) \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}) \alpha_{s-i}}{2(1-\prod_{v=1}^S \beta_v^2)} + \frac{r^2 \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}^2) Q_{s-i}^2}{8(1-\prod_{v=1}^S \beta_v^2)}\right\}. \end{aligned}$$

En plus, si $(\eta_{s+\tau S})$ est normalement distribuée, il est possible d'avoir les résultats explicites suivants

$$\mathbb{E}(\varepsilon_{s+\tau S}^r) = \begin{cases} \frac{(2m)!}{m!2^m} \exp\left\{\frac{2m(1+\prod_{v=1}^S \beta_v) \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}) \alpha_{s-i}}{2(1-\prod_{v=1}^S \beta_v^2)} + \frac{m^2 \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}^2) Q_{s-i}^2}{2(1-\prod_{v=1}^S \beta_v^2)}\right\}, & \text{si } r = 2m, \\ 0, & \text{si } r = 2m + 1, \end{cases}$$

où m est un entier positif. En particulier, pour $r = 2$ et $1 \leq s \leq S$, nous avons

$$\sigma_\varepsilon^{2(s)} = \exp\left\{\frac{2(1+\prod_{v=1}^S \beta_v) \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}) \alpha_{s-i} + \sum_{i=0}^{S-1} (\prod_{j=0}^{i-1} \beta_{s-j}^2) Q_{s-i}^2}{2(1-\prod_{v=1}^S \beta_v^2)}\right\}.$$

Il est bien connu que les distributions de probabilité empiriques de nombreuses séries chronologiques financières (séries de rendements, ou de variations de prix, ou encore du logarithme de ces variations de prix ...) ne correspondent pas généralement à une distribution gaussienne. Les tests classiques de normalité tendent à rejeter nettement l'hypothèse d'une distribution normale. De plus, les densités de probabilité de ces séries présentent des queues épaisses, elles sont dites leptokurtiques. Une mesure de cet effet est obtenue à partir du coefficient de Kurtosis, rapport du moment empirique centré d'ordre 4 et du carré de la variance empirique qui, asymptotiquement, vaut 3 dans le cas d'une loi normale quelconque. Cette valeur sert de référence, et est nettement supérieure pour ces séries.

Pour un processus périodiquement stationnaire (ε_t) solution de (2.2) et sous l'hypothèse que le coefficient d'aplatissement de η_t existe et il est fini (i.e. $\kappa_\eta := \mathbb{E}(\eta_t^4) < \infty$) et la normalité de e_t , nous avons

$$\mathbb{E}(\varepsilon_{s+\tau S}^4) = \mathbb{E}(\eta_{s+\tau S}^4) \exp\{2\mu_x^{(s)} + 2\sigma_x^{2(s)}\},$$

et

$$\mathbb{E}(\varepsilon_{s+\tau S}^2) = \exp\left\{\mu_x^{(s)} + \frac{\sigma_x^{2(s)}}{2}\right\}.$$

Ainsi, le coefficient de Kurtosis de la distribution de $(\varepsilon_{s+\tau S})$ est donné par

$$\begin{aligned} \kappa_\varepsilon^{(s)} &:= \frac{\mathbb{E}(\varepsilon_{s+\tau S}^4)}{[\mathbb{E}(\varepsilon_{s+\tau S}^2)]^2} = \kappa_\eta \times \exp\{\sigma_x^{2(s)}\} \\ &= \kappa_\eta \times \exp\left\{\frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{s-j}^2\right) \left[\left(\alpha_{s-i} + \beta_{s-i} \mu_x^{(s-i-1)} - \mu_x^{(s-i)}\right)^2 + Q_{s-i}^2\right]}{1 - \prod_{v=1}^S \beta_v^2}\right\}. \end{aligned}$$

Il est clair que, comme dans le modèle *SV* classique considéré par Taylor (1982, 1986), les queues de la distribution de (ε_t) sont plus épaisses, lorsque la variance de (x_t) est grande et (ε_t) présente plus d'aplatissement que (η_t) . Par suite, afin d'obtenir les expressions explicites de $\kappa_\varepsilon^{(s)}$, pour $s = \overline{1, S}$, nous avons besoin de supposer une distribution particulière pour (η_t) . Lorsqu'on suppose également que (η_t) est normalement distribuée, nous avons

$$\kappa_\varepsilon^{(s)} = 3 \times \exp\left\{\frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{s-j}^2\right) \left[\left(\alpha_{s-i} + \beta_{s-i} \mu_x^{(s-i-1)} - \mu_x^{(s-i)}\right)^2 + Q_{s-i}^2\right]}{1 - \prod_{v=1}^S \beta_v^2}\right\} > 3.$$

Notons que le minimum est donné par le coefficient d'aplatissement de (η_t) , qui est égal à 3. Par conséquent, les queues de la distribution d'un processus *PAR-SV* sont plus épaisses que celles d'une distribution normale. Ainsi, un modèle *PAR-SV* peut reproduire des queues épaisses habituellement observées dans les séries financières dont la volatilité affiche une structure de corrélation périodique.

Fonctions d'autocovariance et d'autocorrélation d'un modèle *PAR-SV*

La théorie classique des séries temporelles est centrée sur la structure du second ordre des processus. Les processus stationnaires gaussiens sont complètement caractérisés par leur moyenne et leur fonction

d'autocovariance. Pour les processus non gaussiens, il est fréquent, à partir d'une réalisation de longueur n de la série, soit $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, de chercher à estimer la fonction d'autocovariance du processus sous-jacent afin d'avoir une première idée de sa structure de dépendance temporelle. Les praticiens font souvent une utilisation extensive de ce coefficient, du fait de sa facilité d'utilisation. Toutefois, quelques précautions d'usage sont à prendre. Cette étape est préliminaire à toute construction d'un modèle approprié.

Dans cette section, nous étudions la structure de dépendance du processus périodiquement stationnaire du second ordre (ε_t) solution de (2.2), en examinant en premier lieu sa fonction d'autocovariance périodique $\gamma_{\varepsilon,k}^{(s)} := \text{cov}(\varepsilon_{s+\tau S}, \varepsilon_{s+\tau S-k})$.

Proposition 2.1 (Boussaha et Hamdi, 2017) *Soit $\{\varepsilon_t; t \in \mathbb{Z}\}$ un processus périodiquement stationnaire au second ordre solution de (2.2), avec $\prod_{l \geq 0} M_e(\psi_{s,l} Q_{s-l}) < \infty$, pour $s = \overline{1, S}$. Alors $\{\varepsilon_t; t \in \mathbb{Z}\}$ est un processus bruit blanc périodique de variance*

$$\sigma_\varepsilon^{2(s)} := \text{var}(\varepsilon_{s+\tau S}) = \exp \left\{ \frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{s-j} \right) \alpha_{s-i}}{1 - \prod_{v=1}^S \beta_v} \right\} \prod_{l \geq 0} M_e(\psi_{s,l} Q_{s-l}).$$

Preuve. Nous avons pour tout $\tau \in \mathbb{Z}$ et $1 \leq s \leq S$,

$$\begin{aligned} \mathbb{E}(\varepsilon_{s+\tau S}) &= \mathbb{E}(\eta_{s+\tau S}) \mathbb{E} \left(\exp \left\{ \frac{1}{2} x_{s+\tau S} \right\} \right) \\ &= \mathbb{E}(\eta_{s+\tau S}) \exp \left\{ \frac{1}{2} \mu_x^{(s)} \right\} \mathbb{E} \left(\exp \left\{ \frac{1}{2} \sum_{l \geq 0} \psi_{s,l} Q_{s-l} e_{s+\tau S-l} \right\} \right) \\ &= \mathbb{E}(\eta_{s+\tau S}) \exp \left\{ \frac{1}{2} \mu_x^{(s)} \right\} \prod_{l \geq 0} M_e \left(\frac{1}{2} \psi_{s,l} Q_{s-l} \right), \\ &= 0. \end{aligned}$$

D'après le Théorème 2.2, nous avons aussi

$$\sigma_\varepsilon^{2(s)} = \mathbb{E}(\varepsilon_{s+\tau S}^2) = \exp \left\{ \frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{s-j} \right) \alpha_{s-i}}{1 - \prod_{v=1}^S \beta_v} \right\} \prod_{l \geq 0} M_e(\psi_{s,l} Q_{s-l}).$$

De plus, pour tout $(\tau, k) \in \mathbb{Z} \times \mathbb{Z}^*$ et $1 \leq s \leq S$, nous avons $\gamma_{\varepsilon,k}^{(s)} = 0$, comme conséquence immédiate de l'inégalité de Cauchy-Schwarz

$$\left| \gamma_{\varepsilon,k}^{(s)} \right| = |\text{cov}(\varepsilon_{s+\tau S}, \varepsilon_{s+\tau S-k})| \leq \sqrt{\sigma_\varepsilon^{2(s)}} \sqrt{\sigma_\varepsilon^{2(s-k)}}.$$

■

A ce stade, la Proposition 2.1 est cruciale. En effet, même si (ε_t) est un processus bruit blanc périodique, il est possible qu'une dépendance non linéaire existe entre les termes successifs de (ε_t) . Il est bien de rappeler que ce coefficient d'autocorrélation ne mesure que les dépendances linéaires entre les variables, les dépendances non linéaires étant exclues, sauf dans le cas gaussien, où l'absence de

corrélation implique l'indépendance. En plus, d'autres dépendances sur les moments d'ordre supérieur ou égal à 2 peuvent exister ; elles ne sont pas mesurées par ce coefficient. Ainsi, la nature de la dépendance dans un processus de volatilité stochastique périodique défini en (2.2) peut être obtenue en étudiant la structure de covariance du processus des puissances $\{\varepsilon_t^r; t \in \mathbb{Z}, r \in \mathbb{N}^*\}$.

Proposition 2.2 (Boussaha et Hamdi, 2017) *Soit $\{\varepsilon_t; t \in \mathbb{Z}\}$ solution périodiquement stationnaire de (2.2). Sous l'hypothèse de normalité de (η_t) et (e_t) , les fonctions d'autocovariance et d'autocorrélation de période s , $1 \leq s \leq S$, et de retard $k \in \mathbb{Z}$, du processus $\{\varepsilon_t^r; t \in \mathbb{Z}, r \in \mathbb{N}^*\}$ sont données par*

$$\begin{aligned}
\gamma_{\varepsilon^r, k}^{(s)} &:= \text{cov}(\varepsilon_{s+\tau S}^r, \varepsilon_{s+\tau S-k}^r) \\
&= [\mathbb{E}(\eta_{s+\tau S}^r)]^2 \exp \left\{ \sum_{i=0}^{k-1} \left[\left(\prod_{j=0}^{i-1} \beta_{s-j} \right) \frac{r\alpha_{s-i}}{2} + \left(\prod_{j=0}^{i-1} \beta_{s-j}^2 \right) \frac{r^2 Q_{s-i}^2}{8} \right] \right. \\
&\quad \left. + \left(1 + \prod_{j=0}^{k-1} \beta_{s-j} \right) \frac{r\mu_x^{(s-k)}}{2} + \left(1 + \prod_{j=0}^{k-1} \beta_{s-j}^2 \right) \frac{r^2 \sigma_x^{2(s-k)}}{8} \right\} \\
&\quad \times \left[\exp \left\{ \left(\prod_{j=0}^{k-1} \beta_{s-j} \right) \frac{r^2 \sigma_x^{2(s-k)}}{4} \right\} - \exp \{1\} \right], \tag{2.7}
\end{aligned}$$

et

$$\begin{aligned}
\rho_{\varepsilon^r, k}^{(s)} &:= \text{corr}(\varepsilon_{s+\tau S}^r, \varepsilon_{s+\tau S-k}^r) \\
&= \frac{[\mathbb{E}(\eta_{s+\tau S}^r)]^2}{\mathbb{E}(\eta_{s+\tau S}^{2r}) \left[\left(\exp \left\{ \frac{r^2 \sigma_x^{2(s-k)}}{4} \right\} - \exp \{1\} \right) \left(\exp \left\{ \frac{r^2 \sigma_x^{2(s)}}{4} \right\} - \exp \{1\} \right) \right]^{1/2}} \\
&\quad \times \exp \left\{ \sum_{i=0}^{k-1} \left[\left(\prod_{j=0}^{i-1} \beta_{s-j} \right) \frac{r\alpha_{s-i}}{2} + \left(\prod_{j=0}^{i-1} \beta_{s-j}^2 \right) \frac{r^2 Q_{s-i}^2}{8} \right] \right. \\
&\quad \left. + \left(\frac{1}{2} + \prod_{j=0}^{k-1} \beta_{s-j} \right) \frac{r\mu_x^{(s-k)}}{2} + \left(\frac{1}{2} + \prod_{j=0}^{k-1} \beta_{s-j}^2 \right) \frac{r^2 \sigma_x^{2(s-k)}}{8} - \frac{r\mu_x^{(s)}}{4} - \frac{r^2 \sigma_x^{2(s)}}{16} \right\} \\
&\quad \times \left[\exp \left\{ \left(\prod_{j=0}^{k-1} \beta_{s-j} \right) \frac{r^2 \sigma_x^{2(s-k)}}{4} \right\} - \exp \{1\} \right], \tag{2.8}
\end{aligned}$$

où $\mu_x^{(s)}$ et $\sigma_x^{2(s)}$ sont données dans (2.5) et (2.6), respectivement.

Preuve. Nous avons

$$x_{s+\tau S} = \sum_{i=0}^{k-1} \left(\prod_{j=0}^{i-1} \beta_{s-j} \right) (\alpha_{s-i} + Q_{s-i} e_{s+\tau S-i}) + \left(\prod_{j=0}^{k-1} \beta_{s-j} \right) x_{s+\tau S-k},$$

ce qui donne

$$\begin{aligned}
\mathbb{E}(\varepsilon_{s+\tau S}^r \varepsilon_{s+\tau S-k}^r) &= \mathbb{E}(\eta_{s+\tau S}^r) \mathbb{E}(\eta_{s+\tau S-k}^r) \mathbb{E}\left(\exp\left\{\frac{r}{2}(x_{s+\tau S} + x_{s+\tau S-k})\right\}\right) \\
&= \mathbb{E}(\eta_{s+\tau S}^r) \mathbb{E}(\eta_{s+\tau S-k}^r) \mathbb{E}\left(\exp\left\{\frac{r}{2}\sum_{i=0}^{k-1}\left(\prod_{j=0}^{i-1}\beta_{s-j}\right)(\alpha_{s-i} + Q_{s-i}e_{s+\tau S-i})\right.\right. \\
&\quad \left.\left. + \frac{r}{2}\left(1 + \prod_{j=0}^{k-1}\beta_{s-j}\right)x_{s+\tau S-k}\right\}\right) \\
&= [\mathbb{E}(\eta_{s+\tau S}^r)]^2 \times \exp\left\{\sum_{i=0}^{k-1}\left[\left(\prod_{j=0}^{i-1}\beta_{s-j}\right)\frac{r\alpha_{s-i}}{2} + \left(\prod_{j=0}^{i-1}\beta_{s-j}^2\right)\frac{r^2Q_{s-i}^2}{8}\right]\right. \\
&\quad \left. + \left(1 + \prod_{j=0}^{k-1}\beta_{s-j}\right)\frac{r\mu_x^{(s-k)}}{2} + \left(1 + \prod_{j=0}^{k-1}\beta_{s-j}\right)^2\frac{r^2\sigma_x^2(s-k)}{8}\right\}.
\end{aligned}$$

De plus

$$\begin{aligned}
\mathbb{E}(\varepsilon_{s+\tau S}^r) &= \mathbb{E}(\eta_{s+\tau S}^r) \mathbb{E}\left(\exp\left\{\frac{r}{2}x_{s+\tau S}\right\}\right) \\
&= \mathbb{E}(\eta_{s+\tau S}^r) \mathbb{E}\left(\exp\left\{\frac{r}{2}\sum_{i=0}^{k-1}\left(\prod_{j=0}^{i-1}\beta_{s-j}\right)(\alpha_{s-i} + Q_{s-i}e_{s+\tau S-i}) + \frac{r}{2}\left(\prod_{j=0}^{k-1}\beta_{s-j}\right)x_{s+\tau S-k}\right\}\right) \\
&= \exp\left\{\sum_{i=0}^{k-1}\left[\frac{r}{2}\left(\prod_{j=0}^{i-1}\beta_{s-j}\right)\alpha_{s-i} + \frac{r^2}{8}\left(\prod_{j=0}^{i-1}\beta_{s-j}^2\right)Q_{s-i}^2\right]\right. \\
&\quad \left. + \frac{r}{2}\left(\prod_{j=0}^{k-1}\beta_{s-j}\right)\mu_x^{(s-k)} + \frac{r^2}{8}\left(\prod_{j=0}^{k-1}\beta_{s-j}^2\right)\sigma_x^2(s-k)\right\},
\end{aligned}$$

et

$$\begin{aligned}
\mathbb{E}(\varepsilon_{s+\tau S-k}^r) &= \mathbb{E}(\eta_{s+\tau S-k}^r) \mathbb{E}\left(\exp\left\{\frac{r}{2}x_{s+\tau S-k}\right\}\right) \\
&= \mathbb{E}(\eta_{s+\tau S-k}^r) \exp\left\{\frac{r}{2}\mu_x^{(s-k)} + \frac{r^2}{8}\sigma_x^2(s-k)\right\},
\end{aligned}$$

d'où

$$\begin{aligned}
\mathbb{E}(\varepsilon_{s+\tau S}^r) \mathbb{E}(\varepsilon_{s+\tau S-k}^r) &= [\mathbb{E}(\eta_{s+\tau S}^r)]^2 \exp\left\{\sum_{i=0}^{k-1}\left[\left(\prod_{j=0}^{i-1}\beta_{s-j}\right)\frac{r\alpha_{s-i}}{2} + \left(\prod_{j=0}^{i-1}\beta_{s-j}^2\right)\frac{r^2Q_{s-i}^2}{8}\right]\right. \\
&\quad \left. + \left(1 + \prod_{j=0}^{k-1}\beta_{s-j}\right)\frac{r\mu_x^{(s-k)}}{2} + \left(1 + \prod_{j=0}^{k-1}\beta_{s-j}^2\right)\frac{r^2\sigma_x^2(s-k)}{8}\right\}.
\end{aligned}$$

Ainsi, la fonction d'autocovariance de la saison s et du retard $k \in \mathbb{Z}$, du processus $\{\varepsilon_t^r; t \in \mathbb{Z}, r \in \mathbb{N}^*\}$ est donnée par (2.7). En particulier, pour $k = 0$, nous avons

$$\gamma_{\varepsilon^r, 0}^{(s)} = [\mathbb{E}(\eta_{s+\tau S}^r)]^2 \left(\exp\left\{\frac{r^2}{4}\sigma_x^2(s)\right\} - \exp\{1\}\right) \exp\left\{\frac{r}{2}\mu_x^{(s)} + \frac{r^2}{8}\sigma_x^2(s)\right\}.$$

Par conséquent, la fonction d'autocorrélation de la saison s et du retard $k \in \mathbb{Z}$, du processus $\{\varepsilon_t^r; t \in \mathbb{Z}, r \in \mathbb{N}^*\}$ est donnée par (2.8). ■

A partir de cette proposition, il convient de souligner que la séquence des variables aléatoires satisfaisant une représentation *PAR-SV* définie par (2.2) est non corrélée, mais les processus des puissances $\{\varepsilon_t^r; t \in \mathbb{Z}, r \in \mathbb{N}^*\}$ le sont.

2.4 Estimation des paramètres d'un modèle *PAR-SV*

L'estimation des paramètres d'un modèle *PAR-SV* à partir de sa représentation espace d'états est l'objet de cette section. En premier lieu, nous rappelons le problème de l'évaluation de la fonction quasi-vraisemblance d'un modèle *PAR-SV* via la méthode de l'innovation où les innovations empiriques sont obtenues par le filtre de Kalman périodique ou par les équations de Chandrasekhar périodiques (Aknouche et Hamdi, 2007). En second lieu, nous utilisons les méthodes de Monte Carlo Séquentielles (*SMC*), qui offrent un cadre de modélisation informatiquement efficace, et en particulier pour l'estimation des paramètres d'une large gamme de modèles. Nous adaptons plus précisément les algorithmes de filtrage et de lissage particulière appliqués aux modèles espace d'états non périodiques (voir Kim, 2005 ; Kim et Stoffer, 2008) pour l'estimation des paramètres d'un modèle périodique.

2.4.1 Méthode du quasi-maximum de vraisemblance basée sur le filtre de Kalman périodique

Nous allons présenter brièvement la méthode du quasi-maximum de vraisemblance proposée par Aknouche et al. (2007). Considérons le modèle *PAR-SV* défini dans (2.2). Soit $\theta = (\theta'_1, \dots, \theta'_S)'$, où $\theta_s = (\alpha_s, \beta_s, Q_s)'$. Soit également $\widehat{Y}_{s+\tau S|s+\tau S-1}$ le meilleur prédicteur linéaire (au sens des moindres carrés) de $Y_{s+\tau S}$ basé sur $Y_1, Y_2, \dots, Y_{s+\tau S-1}$, (i.e., le vecteur donné par la projection orthogonale du vecteur $Y_{s+\tau S}$ sur le sous-espace engendré par $\{Y_1, Y_2, \dots, Y_{s+\tau S-1}\}$) et $\widehat{u}_{s+\tau S} = Y_{s+\tau S} - \widehat{Y}_{s+\tau S|s+\tau S-1}$ l'innovation empirique à l'instant $s + \tau S$, dont l'erreur moyenne quadratique $\Omega_{s+\tau S} = \mathbb{E}(\widehat{u}_{s+\tau S}^2)$.

L'innovation est aussi donnée par $\widehat{u}_{s+\tau S} = Y_{s+\tau S} - \widehat{x}_{s+\tau S|s+\tau S-1} - d$, où $\widehat{x}_{s+\tau S|s+\tau S-1}$ est le meilleur prédicteur linéaire de $x_{s+\tau S}$ basé sur $Y_1, Y_2, \dots, Y_{s+\tau S-1}$, dont la matrice de covariance de l'erreur de prédiction est

$$P_{s+\tau S|s+\tau S-1} = \mathbb{E} \left[\left(x_{s+\tau S} - \widehat{x}_{s+\tau S|s+\tau S-1} \right)^2 \right].$$

Pour une réalisation donnée¹ $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$ du modèle (2.2), la fonction quasi-vraisemblance du vecteur des paramètres θ , peut être écrite sous la forme d'innovation suivante

$$L(\theta; Y) = (2\pi)^{-\frac{n}{2}} \prod_{s=1}^S \prod_{\tau=0}^{\tau_1-1} \Omega_{s+\tau S}^{-1/2} \exp \left\{ -\frac{1}{2} \sum_{s=1}^S \sum_{\tau=0}^{\tau_1-1} \frac{\widehat{u}_{s+\tau S}^2}{\Omega_{s+\tau S}} \right\},$$

dans laquelle nous avons besoin d'évaluer $\Omega_{s+\tau S}$ et $\widehat{u}_{s+\tau S}$ pour $s = \overline{1, S}$ et $\tau = \overline{0, \tau_1 - 1}$. Pour simplifier et sans perdre de généralité, nous supposons que $n = \tau_1 S$. Ceci peut être achevé en utilisant le filtre de Kalman périodique ou l'algorithme de Chandrasekhar périodique (Aknouche et Hamdi, 2007). Une

¹Dans tous les chapitres, il est supposé que toutes les observations ont été complètement enregistrées.

fois que nous avons exprimé le modèle (2.2) sous forme espace d'états périodique (2.4), la log-quasi-vraisemblance du vecteur des paramètres θ , est donnée par

$$\mathcal{L}(\theta; Y) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{s=1}^S \sum_{\tau=0}^{\tau_1-1} \log \Omega_{s+\tau S} - \frac{1}{2} \sum_{s=1}^S \sum_{\tau=0}^{\tau_1-1} \frac{\widehat{u}_{s+\tau S}^2}{\Omega_{s+\tau S}}. \quad (2.9)$$

Comme dans Aknouche et *al.* (2007), il est possible d'évaluer $\Omega_{s+\tau S}$ et $\widehat{u}_{s+\tau S}$, dont nous avons besoin dans (2.9), à partir des équations du filtre de Kalman périodique données par

$$\begin{cases} \widehat{u}_{s+\tau S} &= Y_{s+\tau S} - \widehat{x}_{s+\tau S|s+\tau S-1} - d, \\ \Omega_{s+\tau S} &= P_{s+\tau S|s+\tau S-1} + \pi^2/2, \\ \widehat{x}_{s+\tau S+1|s+\tau S} &= \alpha_{s+1} + \beta_{s+1} \widehat{x}_{s+\tau S|s+\tau S-1} + K_{s+\tau S} \widehat{u}_{s+\tau S}, \\ K_{s+\tau S} &= \beta_{s+1} P_{s+\tau S|s+\tau S-1} \Omega_{s+\tau S}^{-1}, \\ P_{s+\tau S+1|s+\tau S} &= \beta_{s+1}^2 \left(P_{s+\tau S|s+\tau S-1} - P_{s+\tau S|s+\tau S-1}^2 \Omega_{s+\tau S}^{-1} \right) + Q_{s+1}^2, \end{cases}$$

avec les valeurs de démarrage

$$\widehat{x}_{1|0} = \frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{1-j} \right) \alpha_{1-i}}{1 - \prod_{v=1}^S \beta_v},$$

et

$$P_{1|0} = \frac{\sum_{i=0}^{S-1} \left(\prod_{j=0}^{i-1} \beta_{1-j}^2 \right) \left[\left(\alpha_{1-i} + \beta_{1-i} \mu_x^{(S-i)} - \mu_x^{(1-i)} \right)^2 + Q_{1-i}^2 \right]}{1 - \prod_{v=1}^S \beta_v^2}.$$

En maximisant la quasi log-vraisemblance $\mathcal{L}(\theta; Y)$ par rapport au vecteur θ , nous obtenons $\widehat{\theta}$ l'estimateur du quasi-maximum de vraisemblance (*QML*) de θ . Notons qu'il n'existe pas d'expression explicite de $\widehat{\theta}$. Pour cette raison, nous nous appuyons sur des méthodes d'optimisation numérique. Par suite, ces résultats obtenus nous ont permis de définir une procédure à quatre étapes pour estimer les paramètres du modèle *PAR-SV* par la méthode du quasi-maximum de vraisemblance basée sur le filtre de Kalman périodique. Ces quatre étapes peuvent être récapitulées dans l'algorithme suivant.

Algorithme 2.1 (Méthode *QML* basée sur le filtre de Kalman périodique)

1. Utiliser un estimateur préliminaire¹ du vecteur $\theta^{(0)} = \left(\left(\alpha_1^{(0)}, \beta_1^{(0)}, Q_1^{(0)} \right), \dots, \left(\alpha_S^{(0)}, \beta_S^{(0)}, Q_S^{(0)} \right) \right)'$.
2. Exécuter le filtre de Kalman périodique pour évaluer $\widehat{u}_{s+\tau S}$ et $\Omega_{s+\tau S}$, $s = 1, \dots, S$, $\tau = 0, \dots, \tau_1$.
3. Évaluer la fonction *QML*, $\mathcal{L}(\theta; Y)$. Cette dernière peut être évaluée à partir de la relation (2.9).
4. Utiliser une procédure d'optimisation pour déterminer $\widehat{\theta}$ qui minimise $-2\mathcal{L}(\theta; Y)$.

Il est important de noter que dans des conditions appropriées, l'estimateur *QML* (*QMLE*), $\widehat{\theta}$, qui minimise $-\mathcal{L}(\theta; Y)$, est consistant et asymptotiquement normalement distribué (voir Ljung et Caines, 1979, p. 36). Sa matrice de covariance asymptotique est l'inverse de la matrice d'information asymptotique. Malgré ces propriétés asymptotiques souhaitables et leur large applicabilité, les approximations fournies par le filtre de Kalman périodique et les récurrences de filtre de Chandrasekhar périodique,

¹On peut utiliser la méthode des moments, voir la partie application.

deviennent moins efficaces lorsque la normalité ou la linéarité (ou les deux) ne sont pas vérifiées. Ce qui est le cas ici puisque $\log \varepsilon_t^2$ a été approximée par une distribution normale. En effet, il est bien connu que *QMLE* n'est pas nécessairement le meilleur estimateur lorsque la taille est finie. Guerbyenne et Hamdi (2015) ont montré, à travers plusieurs exemples, les faibles propriétés (à taille finie) de l'estimateur quasi-*ML* dans le cadre des modèles espace d'états à coefficients périodiques.

Afin d'améliorer la performance à taille finie de la méthode *QML*, plusieurs auteurs ont proposé un certain nombre d'approximations basées sur les méthodes de Monte Carlo pour les modèles espace d'états en général, et en particulier pour les modèles à volatilité stochastique. Nous nous concentrons sur les méthodes basées sur l'échantillonnage d'importance.

2.4.2 Méthode du maximum de vraisemblance basée sur l'algorithme *EM* et le filtre particulaire

Considérons la représentation linéarisée (2.4) du modèle *PAR-SV*, que nous supposons périodiquement stationnaire. Soient $\underline{Y} = (x_0, x_1, \dots, x_n, Y_1, Y_2, \dots, Y_n)$ le vecteur des données complètes et $\underline{x} = (x_0, x_1, \dots, x_n)$ le vecteur des données manquantes (les log-volatilités). Pour une réalisation donnée $Y = (Y_1, Y_2, \dots, Y_n)$ du modèle (2.4), et sous l'hypothèse que $x_0 \sim \mathcal{N}(\mu_0, \delta_0^2)$, la fonction de vraisemblance complète du vecteur des paramètres θ , peut être exprimée comme suit

$$\begin{aligned} \mathbf{L}(\theta; \underline{Y}) &= f(x_0) \prod_{t=1}^n f(x_t | \underline{x}_{t-1}) f(Y_t | \underline{Y}_{t-1}, \underline{x}) \\ &= f(x_0) \prod_{t=1}^n f(x_t | x_{t-1}) \prod_{t=1}^n f(Y_t | x_t) \\ &= \frac{1}{\sqrt{2\pi\delta_0^2}} \exp\left(-\frac{(x_0 - \mu_0)^2}{2\delta_0^2}\right) \prod_{t=1}^n \frac{1}{\sqrt{2\pi Q_t}} \exp\left(-\frac{(x_t - \alpha_t - \beta_t x_{t-1})^2}{2Q_t^2}\right) \\ &\quad \times \prod_{t=1}^n \left(C_1 \exp\left\{-\frac{1}{2} \exp(Y_t - x_t)\right\} \times \exp\left\{\frac{1}{2}(Y_t - x_t)\right\} \right), \end{aligned}$$

telle que $\underline{x}_{t-1} = (x_0, x_1, \dots, x_{t-1})$ et $\underline{Y}_{t-1} = (Y_1, Y_2, \dots, Y_{t-1})$. Ainsi, la log-vraisemblance complète est donnée par

$$\begin{aligned} -2 \log \mathbf{L}(\theta; \underline{Y}) &= \log \delta_0^2 + \frac{(x_0 - \mu_0)^2}{\delta_0^2} \\ &\quad + \tau_1 \sum_{s=1}^S \log Q_s^2 + \sum_{s=1}^S \sum_{\tau=0}^{\tau_1-1} \frac{(x_{s+\tau S} - \alpha_s - \beta_s x_{s+\tau S-1})^2}{Q_s^2} \\ &\quad + \sum_{s=1}^S \sum_{\tau=0}^{\tau_1-1} \exp(Y_{s+\tau S} - x_{s+\tau S}) - \sum_{s=1}^S \sum_{\tau=0}^{\tau_1-1} (Y_{s+\tau S} - x_{s+\tau S}) + C_2, \end{aligned}$$

où C_1 et C_2 sont des constantes indépendantes de θ .

Remarquons que la fonction de vraisemblance complète dépend des variables inobservables, à savoir les log-volatilités, donc on ne peut pas l'évaluer directement. Dans ce genre de situation, l'estimation

peut être aisément effectuée par l'algorithme récursif *EM* (Dempster et al., 1977), qui est une approche largement applicable au calcul itératif des estimateurs du maximum de vraisemblance (*MLE*). L'algorithme *EM* fournit une séquence de valeurs $\widehat{\theta}^{(i)}$, $i = 1, 2, \dots$ conçue pour être de plus en plus une bonne approximation du *MLE* via deux étapes : *Expectation-step* (Etape *E*) suivie de *Maximisation-step* (Etape *M*). Pour démarrer la $i^{\text{ème}}$ itération, nous avons l'estimation des paramètres de la dernière itération $\widehat{\theta}^{(i-1)}$, et définissons dans l'étape *E*, la fonction **Q** suivante :

$$\begin{aligned}
\mathbf{Q} \left(\theta, \widehat{\theta}^{(i-1)} \right) &= \mathbb{E} \left[-2 \log \mathbf{L}(\theta; \underline{Y}) \mid Y, \theta^{(i-1)} \right] \\
&= \log \delta_0^2 + \frac{\left(x_0^{(n)} - \mu_0 \right)^2 + P_0^{(n)}}{\delta_0^2} + \sum_{t=1}^n \log Q_t^2 \\
&\quad + \sum_{t=1}^n \frac{\left(x_t^{(n)} - \alpha_t - \beta_t x_{t-1}^{(n)} \right)^2 + P_t^{(n)} + \beta_t^2 P_{t-1}^{(n)} - 2\beta_t P_{t,t-1}^{(n)}}{Q_t^2} \\
&\quad + \sum_{t=1}^n \mathbb{E} \left[\exp(Y_t - x_t) \mid Y, \theta^{(i-1)} \right] - \sum_{t=1}^n \left(Y_t - x_t^{(n)} \right) + C_2 \\
&= \log \delta_0^2 + \frac{\left(x_0^{(n)} - \mu_0 \right)^2 + P_0^{(n)}}{\delta_0^2} + \tau_1 \sum_{s=1}^S \log Q_s^2 \\
&\quad + \sum_{\tau=0}^{\tau_1} \sum_{s=1}^S \frac{\left(x_{s+\tau S}^{(n)} - \alpha_s - \beta_s x_{s+\tau S-1}^{(n)} \right)^2 + P_{s+\tau S}^{(n)} + \beta_s^2 P_{s+\tau S-1}^{(n)} - 2\beta_s P_{s+\tau S, s+\tau S-1}^{(n)}}{Q_s^2} \\
&\quad + \sum_{\tau=0}^{\tau_1} \sum_{s=1}^S \mathbb{E} \left[\exp(Y_{s+\tau S} - x_{s+\tau S}) \mid Y, \theta^{(i-1)} \right] - \sum_{\tau=0}^{\tau_1} \sum_{s=1}^S \left(Y_{s+\tau S} - x_{s+\tau S}^{(n)} \right) + C_2 \tag{2.10}
\end{aligned}$$

où

$$\begin{aligned}
x_{s+\tau S}^{(n)} &= \mathbb{E} \left(x_{s+\tau S} \mid Y, \widehat{\theta}^{(i-1)} \right), \\
P_{s+\tau S}^{(n)} &= \mathbb{E} \left(\left(x_{s+\tau S} - x_{s+\tau S}^{(n)} \right)^2 \mid Y, \widehat{\theta}^{(i-1)} \right),
\end{aligned}$$

et

$$P_{s+\tau S, s+\tau S-1}^{(n)} = \mathbb{E} \left(\left(x_{s+\tau S} - x_{s+\tau S}^{(n)} \right) \left(x_{s+\tau S-1} - x_{s+\tau S-1}^{(n)} \right) \mid Y, \widehat{\theta}^{(i-1)} \right).$$

Avant d'aller à l'étape *M*, nous devons évaluer les quantités $x_t^{(n)}$, $P_t^{(n)}$, $P_{t,t-1}^{(n)}$ et $\mathbb{E} \left[\exp(Y_{s+\tau S} - x_{s+\tau S}) \mid Y, \widehat{\theta}^{(i-1)} \right]$. Notons que ces quantités peuvent être séquentiellement approximées en utilisant les algorithmes de filtrage et de lissage particulière. L'algorithme de l'étape de filtrage récapitule la procédure récursive pour approcher empiriquement la densité a posteriori $p(x_t \mid \mathcal{F}_t)$, où \mathcal{F}_t désigne la σ -algèbre engendrée par l'information disponible jusqu'à l'instant t . L'algorithme proposé ci-dessous vise à généraliser l'algorithme de Kim et Stoffer (2005) au cas du modèle *AR-SV* périodique.

Algorithme 2.2 (Algorithme de filtrage particulière pour un modèle *PAR-SV*)

1. *Initialisation* : pour $j = \overline{1, M}$, générer $f_0^{(j)} \sim p_0(x_0)$ avec les poids initiaux $w_0^{(j)} = 1/M$, où M est le nombre de particules.

2. Pour $s = \overline{1, S}$ et $\tau = \overline{0, \tau_1 - 1}$.

(a) Pour $j = \overline{1, M}$

i. Générer $e_{s+\tau S}^{(j)} \sim \mathcal{N}(0, 1)$.

ii. Calculer $p_{s+\tau S}^{(j)} = \alpha_s + \beta_s f_{s+\tau S-1}^{(j)} + Q_s e_{s+\tau S}^{(j)}$.

iii. Mettre à jour les poids : calculer

$$w_{s+\tau S}^{(j)} = w_{s+\tau S-1}^{(j)} p\left(Y_{s+\tau S} | p_{s+\tau S}^{(j)}\right) \\ \propto w_{s+\tau S-1}^{(j)} \exp\left\{-\frac{1}{2} \exp\left(Y_{s+\tau S} - p_{s+\tau S}^{(j)}\right)\right\} \exp\left\{\frac{1}{2} \left(Y_{s+\tau S} - p_{s+\tau S}^{(j)}\right)\right\}.$$

(b) Pour $j = \overline{1, M}$, normaliser les poids : calculer $\tilde{w}_{s+\tau S}^{(j)} = w_{s+\tau S}^{(j)} / \sum_{j=1}^M w_{s+\tau S}^{(j)}$.

(c) Calculer la mesure de dégénérescence $n^{eff} = 1 / \sum_{j=1}^M \tilde{w}_{s+\tau S}^{(j)2}$.

i. Si $n^{eff} \leq n^T$ (typiquement $n^T = M/2$) : générer $f_{s+\tau S}^{(j)}$ par rééchantillonnage de $p_{s+\tau S}^{(1)}$, ..., $p_{s+\tau S}^{(M)}$, avec les poids $\tilde{w}_{s+\tau S}^{(j)}$, et poser $\tilde{w}_{s+\tau S}^{(j)} = 1/M$.

ii. Sinon $f_{s+\tau S}^{(j)} = p_{s+\tau S}^{(j)}$.

3. Finalement, la séquence des M particules $\left\{f_{s+\tau S}^{(j)}; j = \overline{1, M}\right\}$ est un échantillon de $p(x_{s+\tau S} | \mathcal{F}_{s+\tau S})$, pour $s = \overline{1, S}$ et $\tau = \overline{0, \tau_1 - 1}$.

L'algorithme de lissage particulière suivant, à partir duquel nous prenons en compte les informations disponibles après l'instant t , nous donne des approximations des quantités $x_t^{(n)}$, $P_t^{(n)}$, $P_{t,t-1}^{(n)}$ et $\mathbb{E}\left[\exp(Y_{s+\tau S} - x_{s+\tau S}) \mid Y, \hat{\theta}^{(i-1)}\right]$, qui sont nécessaires dans l'étape E .

Algorithme 2.3 (Algorithme de lissage particulière pour un modèle PAR-SV)

1. Pour $j = \overline{1, M}$, choisir $s_n^{(j)} = f_n^{(i)}$, avec les probabilités $\tilde{w}_n^{(i)}$. Poser $W_n^{(j)} = 1/M$.

2. Pour $j = \overline{1, M}$.

(a) Pour $s = S, \dots, 1$ et $\tau = \tau_1 - 1, \dots, 0$, calculer pour $i = \overline{1, M}$

$$W_{s+\tau S-1|s+\tau S}^{(i)} = \tilde{w}_{s+\tau S-1}^{(i)} \exp\left(\frac{-\left(s_{s+\tau S}^{(j)} - \alpha_s - \beta_s f_{s+\tau S-1}^{(i)}\right)^2}{2Q_s^2}\right).$$

(b) Pour $i = \overline{1, M}$, normaliser les poids lissés : calculer

$$\tilde{W}_{s+\tau S-1|s+\tau S}^{(i)} = W_{s+\tau S-1|s+\tau S}^{(i)} / \sum_{j=1}^M W_{s+\tau S-1|s+\tau S}^{(j)}.$$

(c) Choisir $s_{s+\tau S-1}^{(j)} = f_{s+\tau S-1}^{(i)}$, avec les probabilités $\tilde{W}_{s+\tau S-1|s+\tau S}^{(i)}$.

(d) *Finalement, calculer*

$$\begin{cases} \widehat{x}_{s+\tau S}^{(n)} = \frac{\sum_{j=1}^M s_{s+\tau S}^{(j)}}{M}, \\ \widehat{P}_{s+\tau S}^{(n)} = \frac{\sum_{j=1}^M (s_{s+\tau S}^{(j)} - \widehat{x}_{s+\tau S}^{(n)})^2}{M-1}, \\ \widehat{P}_{s+\tau S, s+\tau S-1}^{(n)} = \frac{\sum_{j=1}^M (s_{s+\tau S}^{(j)} - \widehat{x}_{s+\tau S}^{(n)}) (s_{s+\tau S-1}^{(j)} - \widehat{x}_{s+\tau S-1}^{(n)})}{M}, \\ \mathbb{E} \left[\exp(Y_{s+\tau S} - x_{s+\tau S}) \mid Y, \widehat{\theta}^{(i-1)} \right] = \frac{\sum_{j=1}^M \exp\{Y_{s+\tau S} - s_{s+\tau S}^{(j)}\}}{M}. \end{cases}$$

Après avoir remplacé $x_t^{(n)}$, $P_t^{(n)}$ et $P_{t,t-1}^{(n)}$ par leurs approximations, nous passons donc à l'étape M , où $\widehat{\theta}^{(i)}$ est obtenu par

$$\theta^{(i)} = \arg \max_{\theta} Q \left(\theta, \widehat{\theta}^{(i-1)} \right).$$

Les dérivées premières partielles de (2.10), par rapport aux paramètres inconnus α_s , β_s et Q_s , pour $s = \overline{1, S}$, sont données par

$$\begin{aligned} \frac{\partial \mathbf{Q} \left(\theta, \widehat{\theta}^{(i-1)} \right)}{\partial \beta_s} &= -\frac{2}{Q_s^2} \sum_{\tau=0}^{\tau_1-1} \left\{ x_{s+\tau S-1}^{(n)} x_{s+\tau S}^{(n)} + P_{s+\tau S, s+\tau S-1}^{(n)} \right\} \\ &\quad + \frac{2}{Q_s^2} \sum_{\tau=0}^{\tau_1-1} \left\{ \alpha_s x_{s+\tau S-1}^{(n)} + \beta_s \left(x_{s+\tau S-1}^{(n)2} + P_{s+\tau S-1}^{(n)} \right) \right\}, \\ \frac{\partial \mathbf{Q} \left(\theta, \widehat{\theta}^{(i-1)} \right)}{\partial \alpha_s} &= \frac{-2}{Q_s^2} \sum_{\tau=0}^{\tau_1-1} \left(x_{s+\tau S}^{(n)} - \alpha_s - \beta_s x_{s+\tau S-1}^{(n)} \right), \end{aligned}$$

et

$$\begin{aligned} \frac{\partial \mathbf{Q} \left(\theta, \widehat{\theta}^{(i-1)} \right)}{\partial Q_s} &= \frac{2\tau_1}{Q_s} - \frac{2}{Q_s^3} \sum_{\tau=0}^{\tau_1-1} \left\{ \left(x_{s+\tau S}^{(n)} - \alpha_s - \beta_s x_{s+\tau S-1}^{(n)} \right)^2 \right\} \\ &\quad - \frac{2}{Q_s^3} \sum_{\tau=0}^{\tau_1-1} \left\{ P_{s+\tau S}^{(n)} + \beta_s^2 P_{s+\tau S-1}^{(n)} - 2\beta_s P_{s+\tau S, s+\tau S-1}^{(n)} \right\}. \end{aligned}$$

L'estimateur $\widehat{\theta}^{(i)}$ du paramètre θ peut être obtenu en résolvant le système d'équations suivant :

$$\begin{cases} \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i-1)})}{\partial \beta_s} = 0, \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i-1)})}{\partial \alpha_s} = 0, \quad \text{pour tout } s = \overline{1, S}. \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i-1)})}{\partial Q_s} = 0, \end{cases}$$

Par conséquent, à la $i^{\text{ème}}$ itération, les estimateurs des paramètres α_s , β_s et Q_s sont donnés explicitement par

$$\begin{cases} \widehat{\beta}_s^{(i)} = \frac{\left(\sum_{\tau=0}^{\tau_1-1} x_{s+\tau S}^{(n)} \right) \left(\sum_{\tau=0}^{\tau_1-1} x_{s+\tau S-1}^{(n)} \right) - \tau_1 \sum_{\tau=0}^{\tau_1-1} \left(x_{s+\tau S-1}^{(n)} x_{s+\tau S}^{(n)} + P_{s+\tau S, s+\tau S-1}^{(n)} \right)}{\left(\sum_{\tau=0}^{\tau_1-1} x_{s+\tau S-1}^{(n)} \right)^2 - \tau_1 \sum_{\tau=0}^{\tau_1-1} \left(x_{s+\tau S-1}^{(n)2} + P_{s+\tau S-1}^{(n)} \right)}, \\ \widehat{\alpha}_s^{(i)} = \frac{1}{\tau_1} \left[\sum_{\tau=0}^{\tau_1-1} x_{s+\tau S}^{(n)} - \widehat{\beta}_s^{(i)} \sum_{\tau=0}^{\tau_1-1} x_{s+\tau S-1}^{(n)} \right], \\ \widehat{Q}_s^{(i)} = \left[\frac{1}{\tau_1} \sum_{\tau=0}^{\tau_1-1} \left(x_{s+\tau S}^{(n)} - \widehat{\alpha}_s^{(i)} - \widehat{\beta}_s^{(i)} x_{s+\tau S-1}^{(n)} \right)^2 \right. \\ \left. + \frac{1}{\tau_1} \sum_{\tau=0}^{\tau_1-1} \left(P_{s+\tau S}^{(n)} + \left(\widehat{\beta}_s^{(i)} \right)^2 P_{s+\tau S-1}^{(n)} - 2\widehat{\beta}_s^{(i)} P_{s+\tau S, s+\tau S-1}^{(n)} \right) \right]^{1/2}. \end{cases} \quad (2.11)$$

On répète les étapes E et M jusqu'à un seuil prédéfini où un nombre maximum d'itérations est atteint. Ces résultats nous ont permis de définir une deuxième procédure en deux étapes pour estimer les paramètres d'un modèle $PAR-SV$ par la méthode du maximum de vraisemblance basée sur le filtre particulaire. Ces deux étapes peuvent être récapitulées dans l'algorithme suivant.

Algorithme 2.4 (Méthode ML basée sur l'algorithme EM et le filtre particulaire)

1. Utiliser un estimateur préliminaire $\hat{\theta}^{(0)}$ du vecteur θ .
2. Exécuter l'algorithme EM comme suit. A la $i^{\text{ème}}$ itération :
 - (a) Etape E :
 - i. Appliquer le filtre particulaire (Algorithme 2.2) pour obtenir $\{f_t^{(j)}, j = \overline{1, M}\}$ afin d'approximer $p(x_t | \mathcal{F}_t, \theta^{(i-1)})$.
 - ii. Appliquer le lissage particulaire (Algorithme 2.3) pour obtenir $\{s_t^{(j)}, j = \overline{1, M}\}$ afin d'approximer $p(x_t | \mathcal{F}_n, \theta^{(i-1)})$ et de calculer les quantités $\hat{x}_t^{(n)}, \hat{P}_t^{(n)}$ et $\hat{P}_{t,t-1}^{(n)}$.
 - (b) Etape M : déterminer, pour tout $s = \overline{1, S}$, la valeur de $\hat{\beta}_s^{(i)}, \hat{\alpha}_s^{(i)}$ et $\hat{Q}_s^{(i)}$, à partir de (2.11).

2.5 Application

2.5.1 Étude de simulation

Dans le problème de l'estimation, il est souhaitable d'avoir des estimateurs consistants. Nous espérons que les estimations du modèle $PAR-SV$ approcheront les vrais paramètres lorsque la taille de l'échantillon sera augmentée. Dans le but d'étudier et de comparer les performances des deux procédures d'estimation proposées, nous avons généré plusieurs séries à partir de trois modèles $PAR-SV$ périodiquement stationnaires. Pour ces trois modèles générés, nous avons varié la taille de la série ($n = 500, 750$ et 1000) pour voir l'influence de cette dernière sur la convergence empirique des estimateurs. Dans chaque expérience, le nombre de répliquions indépendantes, T , est égal à 1000 . Notons ici que dans chaque cas nous avons utilisé une séquence de $M = 200$ particules. Les valeurs des paramètres correspondantes ont été choisies afin de satisfaire la condition de stationnarité périodique $\left| \prod_{v=1}^S \beta_v \right| < 1$. Néanmoins, dans le deuxième, modèle les paramètres sont choisis de sorte que l'un des paramètres de persistance (β_3) est supérieur à 1 ; et ce dans le but de tester la sensibilité des deux procédures proposées dans une telle situation.

Les Tableaux 2.1– 2.3 reportent les vraies valeurs des paramètres (TV) de chacun des processus générateurs $PAR-SV$, les moyennes empiriques (MLE pour la méthode ML et $QMLE$ pour la méthode QML) et les écarts-types correspondants ($Sd-MLE$ pour la méthode ML et $Sd-QMLE$ pour la méthode QML) des estimations des 1000 répliquions. Pour démarrer le filtre de Kalman et l'algorithme

EM pour chaque réplcation, nous proposons d'utiliser comme paramètres initiaux, les estimateurs suivants obtenus à partir de la méthode des moments. En effet, nous avons

$$\begin{aligned}
\gamma_{Y,2}^{(t)} &= \mathbb{E} \left(\left(Y_t - \mu_Y^{(t)} \right) \left(Y_{t-2} - \mu_Y^{(t-2)} \right) \right) \\
&= \mathbb{E} \left(\left(\alpha_t + \beta_t x_{t-1} + Q_t e_t + d + u_t - \mu_Y^{(t)} \right) \left(Y_{t-2} - \mu_Y^{(t-2)} \right) \right) \\
&= \mathbb{E} \left(\left(\alpha_t + \beta_t (Y_{t-1} - d - u_{t-1}) + Q_t e_t + d + u_t - \mu_Y^{(t)} \right) \left(Y_{t-2} - \mu_Y^{(t-2)} \right) \right) \\
&= \beta_t \gamma_{Y,1}^{(t-1)} \\
\alpha_t &= \mathbb{E} (\alpha_t) \\
&= \mathbb{E} (Y_t) - d - \mathbb{E} (u_t) - \beta_t (\mathbb{E} (Y_{t-1}) - d - \mathbb{E} (u_{t-1})) - Q_t E (e_t) \\
&= \mu_Y^{(t)} - \beta_t \mu_Y^{(t-1)} + (\beta_t - 1) d
\end{aligned}$$

et

$$\begin{aligned}
\mathbb{E} \left[\left\{ \left(Y_t - \mu_Y^{(t)} \right) - \beta_t \left(Y_{t-1} - \mu_Y^{(t-1)} \right) \right\}^2 \right] \\
&= \mathbb{E} \left[\left\{ \left(x_t + d + u_t - \mu_Y^{(t)} \right) - \beta_t \left(x_{t-1} + d + u_{t-1} - \mu_Y^{(t-1)} \right) \right\}^2 \right] \\
&= \left(\alpha_t + d_t - \mu_Y^{(t)} - \beta_t \left(d - \mu_Y^{(t-1)} \right) \right)^2 + Q_t^2 + (1 + \beta_t^2) \mathbb{E} (u_t^2).
\end{aligned}$$

ce qui donne

$$\left\{ \begin{array}{l}
\widehat{\beta}_s^{(0)} = \frac{\sum_{\tau=0}^{\tau_1-1} (Y_{s+\tau S} - \mu_Y^{(s)}) (Y_{s+\tau S-2} - \mu_Y^{(s-2)})}{\sum_{\tau=0}^{\tau_1-1} (Y_{s+\tau S-1} - \mu_Y^{(s-1)}) (Y_{s+\tau S-2} - \mu_Y^{(s-2)})}, \\
\widehat{\alpha}_s^{(0)} = \mu_Y^{(s)} - \widehat{\beta}_s^{(0)} \mu_Y^{(s-1)} + (\widehat{\beta}_s^{(0)} - 1) d, \\
Q_s^{(0)} = \left\{ \frac{1}{\tau_1} \sum_{\tau=0}^{\tau_1-1} \left(\left(Y_{s+\tau S} - \mu_Y^{(s)} \right) - \widehat{\beta}_s^{(0)} \left(Y_{s+\tau S-1} - \mu_Y^{(s-1)} \right) \right)^2 \right. \\
\left. - \left(\alpha_s - \mu_Y^{(s)} + \beta_s \mu_Y^{(s-1)} + d \left(1 - \widehat{\beta}_s^{(0)} \right) \right)^2 - \left(1 + \widehat{\beta}_s^{(0)2} \right) \mathbb{E} (u_{s+\tau S}^2) \right\}^{1/2},
\end{array} \right.$$

$$\text{où } \mu_Y^{(s)} = \frac{1}{\tau_1} \sum_{\tau=0}^{\tau_1-1} Y_{s+\tau S}.$$

Tableau 2.1. Résultats de simulation d'un modèle *PAR-SV* 2-périodique, avec $n = 500, 750, 1000$ et un nombre de réplifications égal à 1000.

n	TV	MLE	$Sd-MLE$	$QMLE$	$Sd-QMLE$
500	$\alpha_1 = 0.5000$	0.5165	0.1797	0.5005	0.2614
	$\beta_1 = 0.8000$	0.7815	0.0785	0.7985	0.1449
	$Q_1 = 1.0000$	1.0422	0.1190	0.9487	0.3142
	$\alpha_2 = 2.0000$	1.9334	0.1898	2.0308	0.2991
	$\beta_2 = -0.9000$	-0.8754	0.0846	-0.9115	0.1656
	$Q_2 = 1.0000$	1.0608	0.1231	0.9318	0.3733
750	$\alpha_1 = 0.5000$	0.5143	0.1495	0.4963	0.2157
	$\beta_1 = 0.8000$	0.7842	0.0659	0.8044	0.1238
	$Q_1 = 1.0000$	1.0454	0.0954	0.9834	0.2483
	$\alpha_2 = 2.0000$	1.9422	0.1557	2.0115	0.2489
	$\beta_2 = -0.9000$	-0.8743	0.0684	-0.8985	0.1353
	$Q_2 = 1.0000$	1.0568	0.0920	0.9665	0.2719
1000	$\alpha_1 = 0.5000$	0.5105	0.1243	0.5007	0.1829
	$\beta_1 = 0.8000$	0.7895	0.0548	0.8022	0.1036
	$Q_1 = 1.0000$	1.0406	0.0838	0.9846	0.2095
	$\alpha_2 = 2.0000$	1.9432	0.1386	2.0121	0.2094
	$\beta_2 = -0.9000$	-0.8762	0.0598	-0.9012	0.1147
	$Q_2 = 1.0000$	1.0535	0.0800	0.9761	0.2242

Tableau 2.2. Résultats de simulation d'un modèle *PAR-SV* 4-périodique, avec $n = 500, 750, 1000$ et un nombre de réplifications égal à 1000.

n	TV	MLE	$Sd-MLE$	$QMLE$	$Sd-QMLE$
500	$\alpha_1 = 0.5000$	0.5156	0.2394	0.4931	0.3292
	$\beta_1 = 0.8000$	0.7833	0.0879	0.8084	0.1518
	$Q_1 = 1.0000$	1.0967	0.1653	0.8498	0.5078
	$\alpha_2 = -1.0000$	-1.0458	0.2505	-0.9603	0.3502
	$\beta_2 = -0.9000$	-0.8867	0.1001	-0.9115	0.1649
	$Q_2 = 1.0000$	1.0573	0.1567	0.8606	0.4793
	$\alpha_3 = 0.0000$	0.0030	0.4373	0.0068	0.2617
	$\beta_3 = 1.2000$	1.1824	0.1564	1.2160	0.1698
	$Q_3 = 2.0000$	1.9751	0.2782	1.9464	0.4589
	$\alpha_4 = 2.0000$	1.9799	0.3217	1.9853	0.3363
	$\beta_4 = 0.5000$	0.5050	0.0842	0.4966	0.0998
	$Q_4 = 2.0000$	1.9228	0.2178	1.9577	0.2917
750	$\alpha_1 = 0.5000$	0.5148	0.1641	0.5010	0.2286
	$\beta_1 = 0.8000$	0.7851	0.0625	0.8013	0.1060
	$Q_1 = 1.0000$	1.0965	0.1210	0.9501	0.3555
	$\alpha_2 = -1.0000$	-1.0356	0.1750	-0.9826	0.2420
	$\beta_2 = -0.9000$	-0.8887	0.0708	-0.9070	0.1114
	$Q_2 = 1.0000$	1.0556	0.1146	0.9218	0.3459
	$\alpha_3 = 0.0000$	0.0028	0.2970	0.0200	0.2515
	$\beta_3 = 1.2000$	1.1852	0.1033	1.2051	0.1238
	$Q_3 = 2.0000$	1.9923	0.1898	1.9698	0.3010
	$\alpha_4 = 2.0000$	1.9921	0.2269	2.0021	0.2391
	$\beta_4 = 0.5000$	0.5045	0.0588	0.5009	0.0692
	$Q_4 = 2.0000$	1.9503	0.1631	1.9919	0.2113
1000	$\alpha_1 = 0.5000$	0.5138	0.1374	0.5071	0.1853
	$\beta_1 = 0.8000$	0.7852	0.0482	0.7993	0.0842
	$Q_1 = 1.0000$	1.1020	0.0958	0.9425	0.2862
	$\alpha_2 = -1.0000$	-1.0360	0.1422	-0.9711	0.1984
	$\beta_2 = -0.9000$	-0.8896	0.0562	-0.9088	0.0920
	$Q_2 = 1.0000$	1.0528	0.0926	0.9427	0.2777
	$\alpha_3 = 0.0000$	-0.0080	0.2420	0.0037	0.2280
	$\beta_3 = 1.2000$	1.1856	0.0865	1.2042	0.1053
	$Q_3 = 2.0000$	2.0070	0.1594	1.9852	0.2415
	$\alpha_4 = 2.0000$	1.9927	0.1806	1.9975	0.1933
	$\beta_4 = 0.5000$	0.5035	0.0461	0.4999	0.0545
	$Q_4 = 2.0000$	1.9526	0.1338	1.9905	0.1735

Tableau 2.3. Résultats de simulation d'un modèle *PAR-SV* 5-périodique, avec $n = 500, 750, 1000$ et un nombre de réplifications égal à 1000.

n	TV	MLE	$Sd-MLE$	$QMLE$	$Sd-QMLE$
500	$\alpha_1 = -2.0000$	-2.1388	0.3534	-2.0934	0.4852
	$\beta_1 = 0.8000$	0.7692	0.0963	0.7782	0.1414
	$Q_1 = 1.0000$	1.0096	0.1465	0.9059	0.4307
	$\alpha_2 = 0.9000$	0.8569	0.4363	1.0331	0.7166
	$\beta_2 = 1.1000$	1.0776	0.1308	1.1182	0.1853
	$Q_2 = 1.0000$	1.0714	0.1570	0.7733	0.5191
	$\alpha_3 = 7.0000$	6.8217	0.6040	7.0364	0.8498
	$\beta_3 = 1.5000$	1.4517	0.1268	1.5244	0.2095
	$Q_3 = 1.0000$	1.3696	0.1607	0.7531	0.6740
	$\alpha_4 = -4.0000$	-4.0163	0.2618	-4.0194	0.3287
	$\beta_4 = 0.5000$	0.4956	0.0594	0.4970	0.0818
	$Q_4 = 1.0000$	1.0211	0.1673	0.8915	0.4133
	$\alpha_5 = 0.0000$	-0.1076	0.4749	0.0019	0.0053
	$\beta_5 = 0.9000$	0.8736	0.0967	0.9022	0.1085
	$Q_5 = 1.0000$	1.0635	0.1845	0.9951	0.5082
750	$\alpha_1 = -2.0000$	-2.0682	0.3267	-2.0396	0.4361
	$\beta_1 = 0.8000$	0.7902	0.0880	0.8003	0.1408
	$Q_1 = 1.0000$	1.0352	0.1290	0.9376	0.3161
	$\alpha_2 = 0.9000$	0.8409	0.4195	0.9540	0.4774
	$\beta_2 = 1.1000$	1.0397	0.0944	1.1043	0.1813
	$Q_2 = 1.0000$	1.1093	0.1270	0.8994	0.4291
	$\alpha_3 = 7.0000$	6.7965	0.4260	6.9559	0.6814
	$\beta_3 = 1.5000$	1.4401	0.0943	1.4945	0.1652
	$Q_3 = 1.0000$	1.3813	0.1404	0.7133	0.6271
	$\alpha_4 = -4.0000$	-4.0033	0.1586	-4.0031	0.2069
	$\beta_4 = 0.5000$	0.4975	0.0479	0.5023	0.0576
	$Q_4 = 1.0000$	1.0138	0.1268	0.9248	0.3398
	$\alpha_5 = 0.0000$	-0.0918	0.4186	0.0012	0.0050
	$\beta_5 = 0.9000$	0.8767	0.0725	0.8966	0.1059
	$Q_5 = 1.0000$	1.0709	0.1286	0.9426	0.3779
1000	$\alpha_1 = -2.0000$	-2.0656	0.2534	-1.9840	0.3883
	$\beta_1 = 0.8000$	0.7860	0.0617	0.8061	0.1140
	$Q_1 = 1.0000$	1.0277	0.1133	0.9514	0.3046
	$\alpha_2 = 0.9000$	0.8657	0.3447	0.9496	0.4305
	$\beta_2 = 1.1000$	1.0664	0.0928	1.1278	0.1635
	$Q_2 = 1.0000$	1.0853	0.1091	0.8720	0.3879
	$\alpha_3 = 7.0000$	6.8316	0.4285	6.9694	0.5564
	$\beta_3 = 1.5000$	1.4454	0.0821	1.4902	0.1422
	$Q_3 = 1.0000$	1.4083	0.1345	0.9141	0.6084
	$\alpha_4 = -4.0000$	-4.0298	0.1469	-4.0058	0.1919
	$\beta_4 = 0.5000$	0.4993	0.0417	0.4936	0.0510
	$Q_4 = 1.0000$	1.0083	0.1100	0.9426	0.2644
	$\alpha_5 = 0.0000$	-0.0907	0.3562	0.0010	0.0057
	$\beta_5 = 0.9000$	0.8789	0.0647	0.8979	0.0842
	$Q_5 = 1.0000$	1.0718	0.1088	0.9288	0.3470

En premier lieu, on s'intéresse dans cette étude principalement aux biais des estimations et les écarts-types. Les résultats fournissent quelques preuves préliminaires concernant les propriétés à taille finie du $QMLE$ et du MLE dans le cadre des modèles $PAR-SV$. A partir des Tableaux 2.1 – 2.3, on peut voir que même avec une taille relativement petite pour le modèle périodique (taille égale à 500 et 750), les deux procédures d'estimation donnent un bon résultat avec un biais très faible et un écart-type raisonnable, et à chaque fois que la taille de l'échantillon augmente, les biais et les écarts-types empiriques diminuent, ce qui reflète la consistance empirique des estimateurs obtenus par ceux méthodes. On peut voir aussi que les écarts-types empiriques $Sd-QMLE$ sont plus grands que $Sd-MLE$. Par exemple, pour $S = 2$ et $n = 500$ (resp $n = 750$ et $n = 1000$), le $Sd-QMLE$ du paramètre Q_2 est d'environ 303% (respectivement 296% et 280%) plus grand que le $Sd-MLE$. Pour $S = 2$ et $n = 500$ (resp $n = 750$ et $n = 1000$), le $Sd-QMLE$ du paramètre α_1 est d'environ 145% (respectivement 144% et 147%) plus grand que la valeur $Sd-MLE$ correspondante. Notons aussi que tous les paramètres sont significatifs sauf pour les estimations QML de α_1, Q_1 et Q_2 , lorsque $S = 4$ et $n = 500$, et pour α_2 lorsque $S = 5$ et $n = 500$.

En second lieu, il est intéressant d'étudier l'effet de la négligence de la périodicité des paramètres. Plus précisément, lorsque les données sont générées à partir d'un modèle $PAR-SV$ et on les modélise avec un modèle $AR-SV$ non périodique. Pour illustrer comment la négligence de la périodicité peut affecter les résultats de l'estimation et la prédiction de la volatilité, nous avons effectué de nombreuses expériences de Monte Carlo. Nous présentons ici les résultats de 400 séries simulées d'une taille de 3000, générées à partir d'un modèle $AR-SV$ 5-périodique. Les vraies valeurs des paramètres et les paramètres estimés ML du modèle $AR-SV$ périodique et non périodique sont présentées dans le Tableau 2.4. En examinant ces résultats, il est possible de conclure que le produit des coefficients de persistance $\prod_{s=1}^5 \beta_s$ du modèle périodique estimé (0.7872) est assez proche de celui du processus générateur des données (0.8060) alors que le coefficient de persistance pour le modèle non périodique estimé (0.6547) est inférieur à sa valeur réelle.

Tableau 2.4. Estimation des paramètres du modèle $PAR-SV$ et $AR-SV$ non-périodique pour 400 séries simulées du $PAR-SV$ 5-périodique

	Modèle simulé			Modèle $PVAR-SV$ estimé			Modèle $PAR-SV$ estimé		
	α_s	β_s	Q_s	α_s	β_s	Q_s	α	β	Q
$s = 1$	-2.0000	0.9000	0.5000	-1.9658	0.9065	0.5012	-2.6815	0.6547	1.7408
$s = 2$	3.0000	1.5000	0.5000	2.4319	1.4309	0.7782	–	–	–
$s = 3$	0.0000	0.6000	0.5000	0.2479	0.6259	0.5009	–	–	–
$s = 4$	-2.0000	1.3000	0.5000	-2.4439	1.2214	0.6314	–	–	–
$s = 5$	0.5000	0.8000	0.5000	0.6235	0.8128	0.5237	–	–	–

Tableau 2.5. Intervalles de confiance asymptotiques des prévisions à l'horizon 1 aux niveaux $(1 - \alpha) 100\%$ des 400 séries simulées

	$(1 - \alpha) 100\%$							
	50%	60%	70%	80%	90%	95%	99%	
$AR-SV$ périodique	51.06	59.74	67.95	75.91	83.97	88.58	93.79	
$AR-SV$ non-périodique	54.49	60.32	65.66	70.84	76.45	80.12	85.23	

En outre, à partir du Tableau 2.5, il apparaît, d'une part, que les intervalles de confiance asymptotiques

tiques des prévisions déterminés à partir du modèle $PAR-SV$ sont plus proches du niveau de confiance théorique que les intervalles de confiance des prévisions déterminés à partir du modèle $AR-SV$ non périodique. D'autre part, en prenant une seule série générée à partir d'un modèle $PAR-SV$, une observation très intéressante est que, généralement, la volatilité calculée à partir du modèle $PAR-SV$ estimé ressemble davantage à la variabilité de la série générée qu'à la volatilité calculée à partir du modèle $AR-SV$ estimée (voir Figure 2.2). Notons que la volatilité estimée, \hat{h}_t , pour les modèles $AR-SV$ périodique et non périodique est obtenue comme suit

$$\hat{h}_t = \exp(\hat{x}_{t|t-1}),$$

où $\hat{x}_{t|t-1}$ peut être évaluée en utilisant le filtre de Kalman.

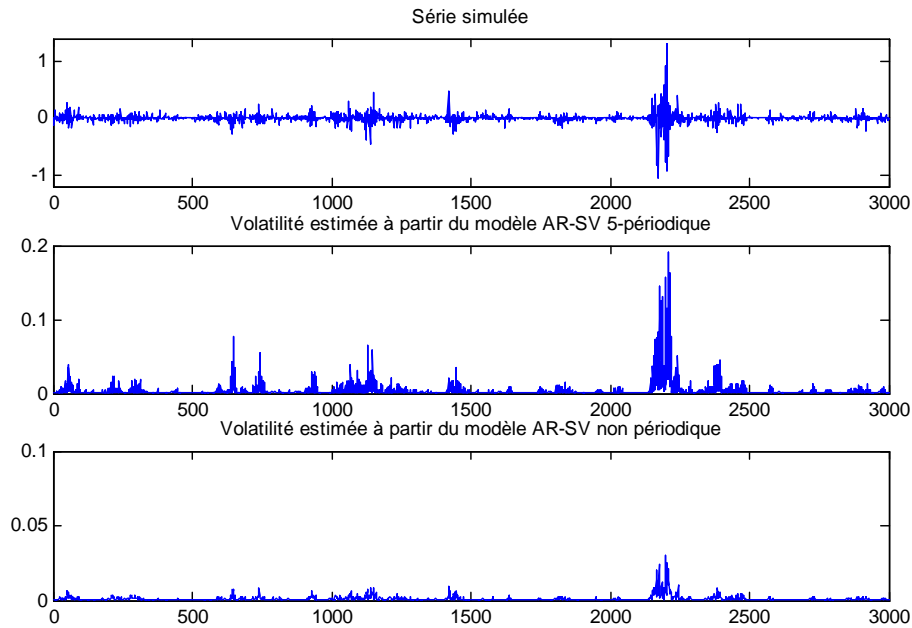


Figure 2.2 Série simulée et volatilités estimées à partir d'un modèle $AR-SV$ périodique et non périodique.

Sur la base de ces observations, nous suggérons l'utilisation d'un modèle périodique à chaque fois que la périodicité est présente.

2.5.2 Application empirique

Pour analyser l'aptitude de la classe de modèles $PAR-SV$ à la modélisation des séries chronologiques avec une structure d'autocorrélation périodique, nous modélisons d'abord la série quotidienne des rendements du taux de change Euro-Dinar Algérien (EUR/DZD) sur la période allant du 3 Janvier 2000 au 29 Septembre 2011 (3056 observations)¹. Cette série a été précédemment analysée par Hamdi et Souam (2013, 2017) et pour laquelle ils ont estimé un $PGARCH(1, 1)$ 5-périodique. Soit y_t le taux de change au comptant EUR/DZD à la date t et $r_t = \log(y_t) - \log(y_{t-1})$ le logarithme du rendement

¹Source de données : Banque d'Algérie.

(également appelé rendement). La série des rendements a une trajectoire compatible avec la stationnarité au second ordre. Il est clair de la représentation graphique de la série r_t , $t = \overline{1, 3055}$, (voir Figure 2.3) que cette dernière est constituée d'oscillations autour de zéro, d'amplitudes très différentes d'une date à l'autre mais de module moyen constant sur des sous-périodes. Notons, cependant, l'extrême volatilité des taux durant la période de la crise financière de 2008. Notons également des sous-périodes de forte agitation des taux (périodes où le marché est plus volatil), suivies de sous-périodes beaucoup plus calmes (périodes de faible volatilité). Le phénomène de regroupement de volatilité apparaît donc clairement.

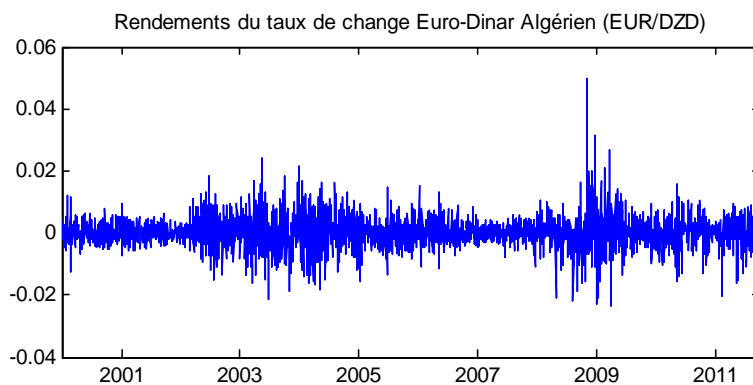


Figure 2.3 Série des rendements du taux de change quotidien Euro-Dinar Algérien (EUR/DZD)

Tableau 2.6. Statistiques descriptives des rendements EUR/DZD de chaque jour de la semaine ainsi que de toute la période d'étude

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Tous les jours
# d'obs	611	611	612	610	611	3055
Moyenne	-0.000141	0.000371	0.000077	0.000216	0.000068	0.000118
Médiane	-0.000078	0.000231	0.000211	0.000070	0.000205	0.000123
Maximum	0.018484	0.031299	0.049695	0.021546	0.024110	0.049695
Minimum	-0.022996	-0.016428	-0.019452	-0.022016	-0.023324	-0.023324
Ecart-type	0.004582	0.004779	0.005205	0.004824	0.005741	0.005043
Skewness	-0.260025	0.571090	1.545905	-0.073088	-0.105271	0.353574
Kurtosis	5.936013	7.146701	18.89839	6.396381	4.875570	8.967818
Jarque-Bera	226.3405	470.9716	6689.1120	293.7345	90.6849	4597.1270

Le Tableau 2.6 présente quelques statistiques des rendements du taux de change EUR/DZD de toute la période d'étude ainsi que les rendements de chaque jour de la semaine. Pour la période analysée, les statistiques descriptives par jour enregistrent le plus grand rendement moyen (0,000371) pour la journée de mardi, alors que le plus petit rendement moyen (-0,000141) apparaît le lundi. Le tableau 2.6 indique également que l'écart-type le plus bas est de 0,004582 pour le lundi, alors que l'écart-type le plus élevé de 0,005741 est observé pour le vendredi. L'asymétrie pour lundi, jeudi et vendredi est négative, alors qu'elle est positive pour mardi et mercredi. De plus, le coefficient de kurtosis de chaque jour de la semaine est beaucoup plus grand que 3, avec la plus petite valeur (4,875570) correspondant au vendredi et la plus grande valeur (18,89839) au mercredi. De plus, les statistiques de Jarque-Bera sont très significatives et soulignent la non-normalité des distributions des rendements EUR/DZD (pour chaque jour de la semaine ainsi que pour toute la période d'étude).

D'autre part, à partir des autocorrélations périodiques d'ordre 1 des carrés et des valeurs absolues des rendements EUR/DZD (voir Tableau 2.7), il est évident que la fonction d'autocorrélation varie sensiblement au cours des jours de la semaine. Cela n'est pas surprenant car il est généralement observé, dans les séries chronologiques financières, que les corrélations entre les rendements quotidiens et la volatilité affichent, certains effets du jour de la semaine (voir par exemple Franses et Paap, 2000 ; Hamdi et Souam, 2017). Nous pouvons donc dire que chaque jour de la semaine constitue un régime différent et c'est pourquoi nous considérons la périodicité comme un enjeu important dans notre modélisation de la série des rendements EUR/DZD .

Tableau 2.7. Autocorrélations périodiques d'ordre 1 des carrés et des valeurs absolues des rendements EUR/DZD

	Lundi	Mardi	Mercredi	Jeudi	Vendredi
valeurs absolues des rendements	0.1812	0.1939	0.3216	0.2252	0.1260
carrés des rendements	0.1142	0.2424	0.3041	0.1614	0.0547

Tableau 2.8. Résultats des estimations des modèles $PAR-SV$ pour la série des rendements EUR/DZD

		Modèle 1			Modèle 2		
		α_s	β_s	Q_s	α_s	β_s	Q_s
Lundi	($s = 1$)	-2.1183	0.8485	0.1618	-2.1894	0.8360	0.0373
Mardi	($s = 2$)	0.9161	1.0610	0.1988	0.7932	1.0562	0.0003
Mercredi	($s = 3$)	7.8441	1.7319	0.4293	5.8887	1.5510	0.00003
Jeudi	($s = 4$)	-4.8566	0.5495	0.2334	-2.8648	0.7341	0.1534
Vendredi	($s = 5$)	-0.0947	0.9657	0.2131	-0.2017	0.9518	0.1702

Dans le Tableau 2.8, nous présentons les résultats d'estimation des paramètres du modèle défini dans (2.2). Les paramètres du modèles $PAR-SV$ ont été estimés en utilisant les deux méthodes ML et QML présentées dans ce chapitre. A partir de ces résultats, nous pouvons remarquer que les deux modèles estimés sont périodiquement stationnaires. En effet, le produit des coefficients de persistance $\prod_{s=1}^5 \beta_s$ est égal à 0.8273 (resp. 0.9570) pour le modèle estimé par la méthode ML (dorénavant Modèle 1) (resp. pour le modèle estimé par méthode QML (dorénavant Modèle 2)), ce qui est en harmonie avec la pratique.

Il est intéressant de relever que les coefficients de persistance estimés pour les modèles $PAR-SV$ pour les journées de Mardi et Mercredi sont supérieurs à 1, contrairement à ceux de Lundi, Jeudi et Vendredi. Ainsi, les chocs de la volatilité qui se produisent le lundi, le jeudi et le vendredi semblent moins informés sur la volatilité future et tendent à disparaître à un rythme plus rapide que les chocs qui se produisent le mardi et le mercredi.

De plus, à partir de la Figure 2.4, il apparaît clairement que pour les deux modèles $PAR-SV$ estimés, le regroupement de la volatilité a été capturé et la volatilité estimée (la volatilité prédite) à partir du modèle $PAR-SV$ ressemble à la variabilité de la série des rendements du taux de change EUR/DZD plus que la volatilité estimée à partir du modèle $PGARCH$. En effet, l'augmentation de la volatilité en 2002, le pic à la fin de 2008 et le ralentissement entre décembre 2006 et mars 2008 peuvent être clairement observés à partir du graphe de la volatilité estimée. Ainsi, le modèle $PAR-SV$ est plus précis et plus approprié pour représenter les caractéristiques de cette série.

Tableau 2.9. Intervalles de prévision asymptotiques à horizon 1, à $(1 - \alpha) 100\%$, pour les rendements *EUR/DZD*.

	$(1 - \alpha) 100\%$						
	50%	60%	70%	80%	90%	95%	99%
<i>PGARCH</i>	38.33	46.28	54.37	63.67	75.19	82.23	91.98
<i>PAR-SV</i> (Modèle 1)	51.49	61.01	70.11	79.15	88.18	92.54	96.69
<i>PAR-SV</i> (Modèle 2)	51.26	60.69	69.82	78.85	88.51	92.80	97.25

Notons aussi qu'à partir du Tableau 2.9, il s'avère que les intervalles de prévision asymptotiques établis à partir de la modélisation *PAR-SV*, sont plus proches du niveau de confiance théorique que, les intervalles de prévision asymptotiques établis à partir de la modélisation *PGARCH*. Ceux-ci montrent que le modèle *PAR-SV* estimé pour la série des rendements quotidiens *EUR/DZD* semble plus précis et plus performant que le modèle *PGARCH*.

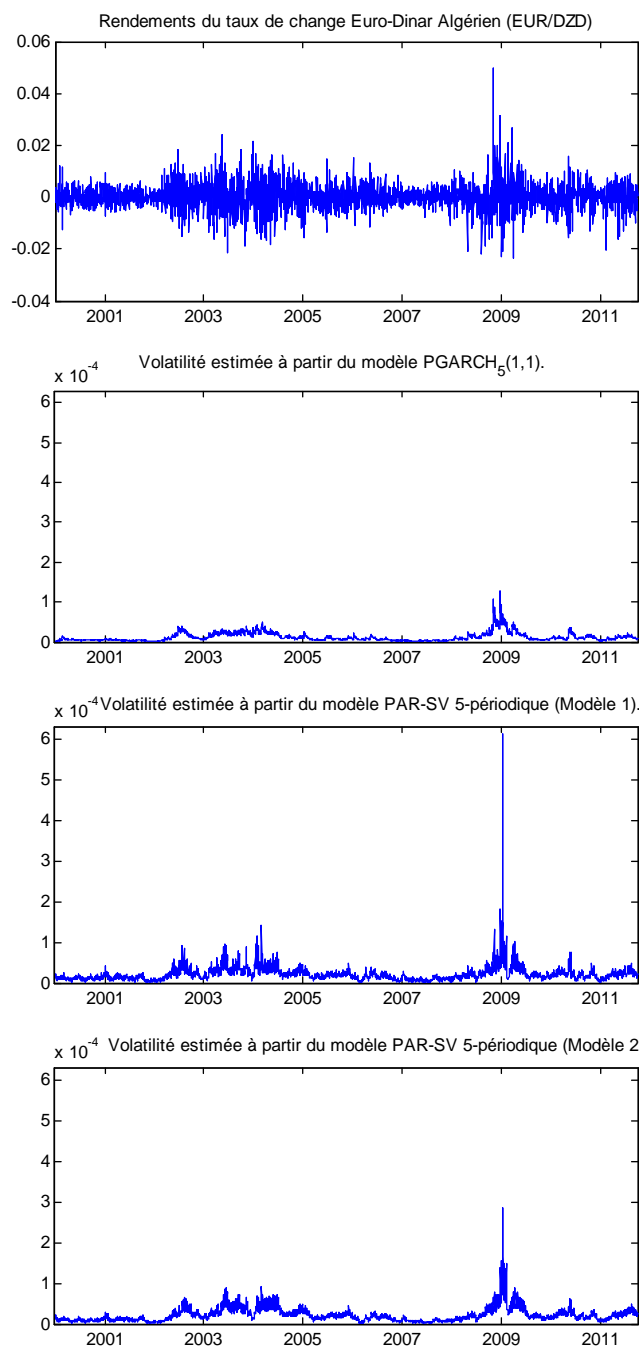


Figure 2.4 Volatilités estimées à partir des modèles 5-périodique *PGARCH* et *PAR-SV*.

Comme deuxième exemple d'application, nous considérons maintenant les taux de change du Dollar Américain par rapport au Dinar Algérien (*USD/DZD*) sur la période allant du 3 Janvier 2000 au 29 Septembre 2011, soit 3056 observations¹. Notons que pour cette série, nous avons limité notre comparaison aux modèles *AR-SV* périodique² et non-périodique ainsi que le modèle *PGARCH* proposé par Hamdi et Souam (2013), ceci afin d'étudier l'effet de la négligence de la périodicité dans la modélisation. La Figure 2.5 et les Tableaux 2.10 – 2.13 donnent les résultats de notre analyse où la comparaison entre les différents modèles estimés donne exactement les mêmes résultats que pour le taux de change *EUR/DZD*.

¹Source de données : Banque d'Algérie.

²Le modèle *AR-SV* périodique est estimé par la méthode *ML*.

Tableau 2.10. Statistiques descriptives des rendements USD/DZD de chaque jour de la semaine ainsi que de toute la période d'étude

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	Tous les jours
# d'obs	611	611	612	610	611	3055
Moyenne	0.012527	-0.016226	0.009650	-0.004363	0.006232	0.001568
Médiane	0.002300	-0.006000	-0.001050	0.000000	0.000000	-0.000500
Maximum	0.740000	1.135200	1.253100	2.211000	1.580300	2.211000
Minimum	-0.978000	-0.717000	-0.856400	-1.180600	-1.389900	-1.389900
Ecart-type	0.196515	0.208196	0.222460	0.236257	0.246081	0.222733
Skewness	0.021923	0.564417	0.826466	0.836611	0.159888	0.501822
Kurtosis	6.244755	7.108851	8.676924	19.18609	9.668562	11.26270
Jarque-Bera	268.0854	462.2449	891.4711	6730.061	1134.728	8818.704

Tableau 2.11. Autocorrélations périodiques d'ordre 1 des carrés et des valeurs absolues des rendements USD/DZD

	Monday	Tuesday	Wednesday	Thursday	Friday
valeurs absolues des rendements	0.3055	0.3620	0.4322	0.3682	0.3217
carrés des rendements	0.1775	0.2493	0.2821	0.2744	0.1359

Tableau 2.12. Résultats des estimations des modèles $AR-SV$ pour la série des rendements USD/DZD

		$AR-SV$ périodique			$AR-SV$ non-périodique		
		α_s	β_s	Q_s	α	β	Q
Lundi	($s = 1$)	0.6106	1.0788	0.5724	-0.3914	0.9697	0.5088
Mardi	($s = 2$)	-0.0691	0.9909	0.4877	-	-	-
Mercredi	($s = 3$)	-0.3029	0.9745	0.6120	-	-	-
Jeudi	($s = 4$)	0.2396	1.0179	0.5035	-	-	-
Vendredi	($s = 5$)	-2.0837	0.8159	0.2328	-	-	-

Tableau 2.13. Intervalles de prévision asymptotiques à horizon 1, à $(1 - \alpha) 100\%$, pour les rendements USD/DZD .

	$(1 - \alpha) 100\%$							
	50%	60%	70%	80%	90%	95%	99%	
$PGARCH$	52.58	60.25	68.68	77.08	85.63	89.95	95.05	
$AR-SV$ périodique	52.62	60.54	68.76	77.18	85.76	90.01	95.09	
$AR-SV$ non-périodique	52.42	60.18	68.89	77.05	85.53	89.91	94.63	

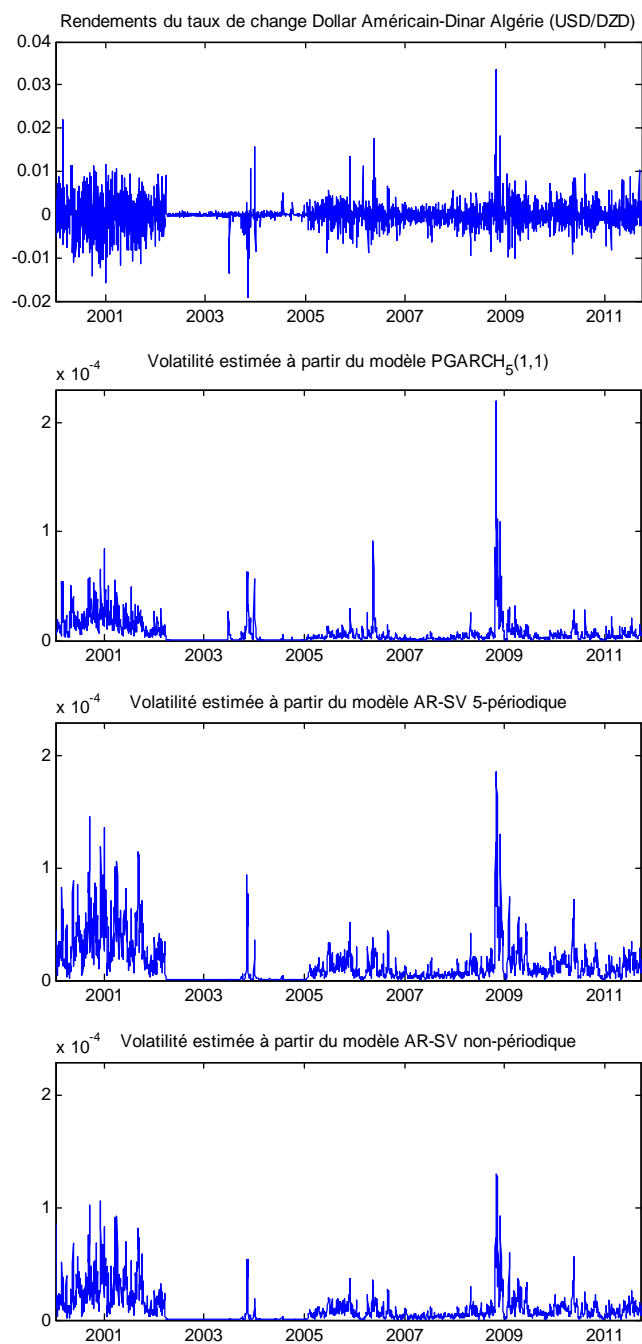


Figure 2.5 Volatilités estimées à partir des modèles 5-périodique *PGARCH* et *PAR-SV*.

Chapitre 3

Modèles à Volatilité Stochastique Multivariés à Coefficients Périodiques

3.1 Introduction

Outre le phénomène de la volatilité et de son regroupement, d'autres faits stylisés importants sont associés aux séries de rendements financiers. On peut citer l'excès de Kurtosis, l'asymétrie, les queues épaisses des distributions et la persistance, etc. En pratique, il s'avère qu'une grande partie de la littérature est basée sur des modèles univariés. Certains faits stylisés ne peuvent, toutefois, pas être capturés par une description univariée. C'est, par exemple, le cas de l'effet de covariation, c'est-à-dire l'étude des relations entre les volatilités et les co-volatilités de plusieurs marchés. D'une part, une grande partie de la prise de décision financière, comme l'optimisation du portefeuille, la répartition des actifs, la gestion des risques et la tarification des actifs, doit clairement prendre en compte des corrélations. D'autre part, il est désormais bien établi, que la volatilité financière des différents actifs et marchés, évolue au fil du temps. Les grands changements d'un actif vont de pair avec de grands changements d'un autre actif. Cette caractéristique joue un rôle crucial dans la modélisation des séries chronologiques financières. Par conséquent, travailler avec un cadre de modélisation multivariée conduit à des modèles empiriques plus pertinents qu'avec des modèles univariés séparés. Certains modèles à volatilité stochastique multivariés sont récemment devenus une préoccupation majeure dans l'étude de la structure de corrélation des séries économiques multivariées en général et des séries temporelles financières multivariées en particulier. En effet, diverses extensions du modèle de base multivarié SV (MSV), introduit par Harvey et al. (1994), ont été proposées dans la littérature. On peut citer, à titre d'exemple, Daniélsson (1998), Jungbacker et Koopman (2006), Smith et Pitts (2006), Chan et al. (2006), McAleer, (2005), Asai et McAleer (2009a, b), etc. Asai et al. (2006) proposent une revue détaillée de cette littérature.

Notons que ces modèles ont été largement utilisés dans la théorie de la finance moderne. De plus, une volumineuse littérature empirique s'est développée. L'extension au cas multivarié n'est pas immédiate puisque le chemin latent des volatilités est converti en un chemin de vecteurs avec des entrées, les

éléments d'une matrice de covariance. Yu et Meyer (2006) ont examiné, en détail, un large éventail de modèles *MSV* dominants, apparus dans la littérature multivariée de la modélisation de la volatilité. Ils ont discuté la spécification et le problème de l'estimation de quelques modèles *MSV*.

Il est important de noter que tous les modèles proposés traitent les paramètres de la volatilité comme constants dans le temps. Il ne semble pas y avoir de formulation capable d'expliquer de manière adéquate les séries chronologiques multivariées dont la structure évolue au fil du temps. C'est en particulier le cas des séries temporelles avec une structure d'autocorrélation périodique (voir chapitre 2).

La contribution dans ce chapitre est double. Nous proposons, dans un premier temps, un modèle, appelé modèle à volatilité stochastique autorégressif multivarié à coefficients périodiques (*PVAR-SV*), afin de capturer la périodicité dans la variance conditionnelle stochastique, en utilisant une spécification qui inclut non seulement la structure de dépendance de chaque composante de la volatilité, mais aussi l'interdépendance entre différentes composantes (la causalité de Granger dans la volatilité). Dans un second temps, nous présentons une application empirique qui montre que notre modélisation donne des résultats très précis et a une bonne performance de prévision de la volatilité.

Ce chapitre est organisé comme suit. La Section 3.2 décrit la classe des modèles *PVAR-SV*. Dans la Section 3.3, nous présentons quelques propriétés probabilistes de cette classe de modèles. Dans la Section 3.4, nous adaptons les mêmes méthodes d'estimation proposées dans le chapitre précédent, à savoir la méthode basée sur le filtre de Kalman périodique et celle basée sur le filtrage particulière combiné avec l'algorithme *EM*. Enfin, dans la Section 3.5, nous appliquons notre nouveau modèle à un ensemble de trois séries de taux de change et au prix du pétrole, à savoir le dollar américain/dinar algérien (*USD/DZD*), l'euro/dinar algérien (*EUR/DZD*), l'euro/dollar américain (*EUR/USD*) et le prix quotidien du pétrole Saharan Blend.

3.2 Modèle à volatilité stochastique autorégressif multivarié à coefficients périodiques

Soient $\boldsymbol{\varepsilon}_t = (\varepsilon_{t,1}, \dots, \varepsilon_{t,m})'$ m -vecteur des rendements de stock à l'instant t et $X_t = (X_{t,1}, \dots, X_{t,m})'$ le vecteur correspondant à la *log-volatilité*. Le modèle à volatilité stochastique autorégressif multivarié à coefficients périodiques (*PVAR-SV*), peut être défini par

$$\begin{cases} \boldsymbol{\varepsilon}_t = V_t^{1/2} \boldsymbol{\eta}_t, \\ X_t = \boldsymbol{\alpha}_t + \Phi_t X_{t-1} + \mathbf{e}_t, \end{cases} \quad t \in \mathbb{Z}, \quad (3.1)$$

où $V_t = \text{diag}(\exp X_{t,1}, \dots, \exp X_{t,m}) = \text{diag}(\exp X_t)$, $\Phi_t = (\phi_{i,j}^{(t)})_{i,j=1,m}$ est une matrice triangulaire inférieure¹, i.e. $\phi_{i,j}^{(t)} = 0$ si $i < j$, et $\boldsymbol{\alpha}_t = (\alpha_{t,1}, \alpha_{t,2}, \dots, \alpha_{t,m})'$. Les deux processus d'innovation m -varié ($\boldsymbol{\eta}_t$) et (\mathbf{e}_t) sont supposés *i.i.d.* et mutuellement indépendants de moyenne nulle avec

$$\mathbb{E}(\boldsymbol{\eta}_t \boldsymbol{\eta}_t') = \Sigma_{\boldsymbol{\eta}}^{(t)}, \quad \mathbb{E}(\mathbf{e}_t \mathbf{e}_t') = \Sigma_{\mathbf{e}}^{(t)} \quad \text{et} \quad \mathbb{E}(\boldsymbol{\eta}_t \mathbf{e}_t') = \mathbf{0}_m,$$

¹On peut considérer (3.1) comme une extension de la formulation Granger causality-*MSV* (*GC-MSV*) proposée par Yu et Meyer (2006).

où $\Sigma_{\boldsymbol{\eta}}^{(t)}$ et $\Sigma_{\mathbf{e}}^{(t)}$ sont des matrices $m \times m$ définies positives. Les paramètres $\boldsymbol{\alpha}_t$, Φ_t , $\Sigma_{\boldsymbol{\eta}}^{(t)}$ et $\Sigma_{\mathbf{e}}^{(t)}$ sont périodiques en t , de période S .

Pour ressortir la périodicité, comme nous avons fait dans le cas univarié, soit $t = s + \tau S$, pour $\tau \in \mathbb{Z}$ et $1 \leq s \leq S$. Alors, le modèle (3.1) peut être réécrit sous la forme suivante

$$\begin{cases} \boldsymbol{\varepsilon}_{s+\tau S} = \text{diag} \left(\exp \left\{ \frac{1}{2} X_{s+\tau S} \right\} \right) \boldsymbol{\eta}_{s+\tau S}, & \tau \in \mathbb{Z}, 1 \leq s \leq S. \\ X_{s+\tau S} = \boldsymbol{\alpha}_s + \Phi_s X_{s+\tau S-1} + \mathbf{e}_{s+\tau S}, \end{cases} \quad (3.2)$$

L'objet de ce qui suit est de trouver des conditions de stationnarité périodique.

3.3 Étude Probabiliste du modèle $PVAR-SV$

3.3.1 Stationnarité périodique et calcul des moments

Examinons maintenant quelques propriétés fondamentales du processus $\{\boldsymbol{\varepsilon}_t; t \in \mathbb{Z}\}$, satisfaisant une représentation $PVAR-SV$ définie dans (3.1). Plus précisément, nous examinons les conditions d'existence d'une solution *p.s.s.* et périodiquement ergodique (telles qu'elles sont définies, par exemple, dans Aknouche et Bibi, 2009).

Il est clair, à partir de la forme multiplicative de la deuxième équation du modèle (3.2), que le processus $\{X_t; t \in \mathbb{Z}\}$ est *p.s.s.* et périodiquement ergodique si et seulement si le rayon spectral de la matrice $\prod_{v=1}^S \Phi_v$ est inférieur à 1 (voir par exemple Aknouche et Bibi (2009) pour le cas d'un modèle $PGARCH$).

Soient le symbole \otimes représentant le produit de Kronecker, $\text{vec}(A)$ l'opérateur usuel de la transformation de la matrice A en vecteur et $A^{\otimes r} := A \otimes A \otimes \dots \otimes A$ le produit de Kronecker de r fois la matrice A , où $r \in \mathbb{N}$ (par convention $A^{\otimes 0} = \mathbf{I}$ et $A^{\otimes 1} = A$). Notons par $\rho(A)$ le rayon spectral de la matrice carrée A .

Il est bien connu que si $\rho\left(\prod_{v=1}^S \Phi_v\right) < 1$, le processus $\{X_t; t \in \mathbb{Z}\}$ est un processus *p.s.s.* d'une moyenne périodique inconditionnelle

$$\mu_X^{(s)} := \mathbb{E}(X_{s+\tau S}) = \left(\mathbf{I}_m - \prod_{v=0}^{S-1} \Phi_{S-v} \right)^{-1} \sum_{v=0}^{S-1} \left(\prod_{i=0}^{v-1} \Phi_{S-i} \right) \boldsymbol{\alpha}_{s-v}, \text{ pour } \tau \in \mathbb{Z} \text{ et } s \in \{1, 2, \dots, S\},$$

et d'un moment d'ordre 2 périodique inconditionnel

$$\begin{aligned} \text{vec} \left(\Sigma_X^{(s)} \right) &:= \mathbb{E} \left(X_{s+\tau S}^{\otimes 2} \right) \\ &= \left(\mathbf{I}_{m^2} - \left(\prod_{v=0}^{S-1} \Phi_{S-v} \right)^{\otimes 2} \right)^{-1} \sum_{v=0}^{S-1} \left(\prod_{i=0}^{v-1} \Phi_{S-i} \right)^{\otimes 2} \text{vec} \left(\Sigma_{\mathbf{e}}^{(s-v)} \right), \text{ pour } \tau \in \mathbb{Z} \text{ et } s \in \{1, 2, \dots, S\}. \end{aligned}$$

Si $\rho\left(\prod_{v=1}^S \Phi_v\right) < 1$, le processus X_t est donné comme un modèle causal VAR périodique d'ordre 1

et peut être représenté comme suit

$$X_{s+\tau S} = \mu_X^{(s)} + \sum_{l \geq 0} \left(\prod_{i=0}^{l-1} \Phi_{s-i} \right) \mathbf{e}_{s+\tau S-l}.$$

Puisque $\boldsymbol{\varepsilon}_t$ est le produit de deux processus *p.s.s.*, il doit être également *p.s.s.* Ainsi, la condition nécessaire et suffisante assurant la stationnarité périodique stricte de $\boldsymbol{\varepsilon}_t$, est la même condition nécessaire et suffisante assurant la stationnarité périodique stricte d'un processus vectoriel autorégressif d'ordre 1 (*PVAR*(1)). Ceci qui est résumé dans le théorème suivant.

Théorème 3.3 (Boussaha et al., 2017) *Le modèle (3.1) admet une unique solution p.s.s. et périodiquement ergodique non anticipative donnée par*

$$\boldsymbol{\varepsilon}_{s+\tau S} = \text{diag} \left(\exp \left\{ \frac{1}{2} \left(\mu_X^{(s)} + \sum_{l \geq 0} \left(\prod_{i=0}^{l-1} \Phi_{s-i} \right) \mathbf{e}_{s+\tau S-l} \right) \right\} \right) \boldsymbol{\eta}_{s+\tau S}, \quad \tau \in \mathbb{Z} \text{ et } s \in \{1, 2, \dots, S\}, \quad (3.3)$$

où la série définie dans (3.3) converge presque sûrement, si et seulement si

$$\rho \left(\prod_{v=1}^S \Phi_v \right) < 1. \quad (3.4)$$

Nous nous intéressons maintenant au calcul explicite de la variance, du coefficient d'asymétrie et d'aplatissement du modèle *PVAR-SV* défini dans (3.1). Pour cela, nous supposons que la condition (3.4) est vérifiée et que les deux vecteurs $\boldsymbol{\eta}_t$ et \mathbf{e}_t sont gaussiens. Dans un premier temps, nous donnons des résultats très généraux. Dans un deuxième temps, nous considérons la normalité afin d'obtenir une formule plus explicite pour les moments. Nous avons donc

$$\begin{aligned} \mathbb{E} (\boldsymbol{\varepsilon}_t^{\otimes r}) &= \mathbb{E} \left(\left(V_t^{1/2} \boldsymbol{\eta}_t \right)^{\otimes r} \right) \\ &= \mathbb{E} \left[\text{diag} \left\{ \left(\exp \frac{1}{2} X_t \right)^{\otimes r} \right\} \times \boldsymbol{\eta}_t^{\otimes r} \right] \\ &= \text{diag} \left\{ \mathbb{E} \left[\left(\exp \frac{1}{2} X_t \right)^{\otimes r} \right] \right\} \mathbb{E} [\boldsymbol{\eta}_t^{\otimes r}] \\ &= \text{diag} \left\{ \mathbb{E} [Y_t^{(r)}] \right\} \mathbb{E} [\boldsymbol{\eta}_t^{\otimes r}], \end{aligned}$$

où

$$\begin{aligned} Y_t^{(r)} &= \left(\exp \frac{1}{2} X_t \right)^{\otimes r} = \left(Y_{t,1}^{(r)}, Y_{t,2}^{(r)}, \dots, Y_{t,m^r}^{(r)} \right), \\ Y_{t,i}^{(r)} &= \exp \left(\frac{1}{2} \sum_{n=0}^{r-1} X_{t,k_{i,n}} \right), \quad \forall i = \overline{1, m^r}, \end{aligned}$$

avec

$$i - k_{i,0} = (k_{i,1} - 1) m + \dots + (k_{i,r-1} - 1) m^{r-1},$$

et

$$k_{i,0} = \begin{cases} m & \text{si } i \equiv 0 \pmod{m}, \\ l & \text{si } i \equiv l \pmod{m} \end{cases}$$

A partir de cette dernière équation, nous avons la relation suivante

$$\mathbb{E} \left(Y_{t,i}^{(r)} \right) = \mathbb{E} \left(\exp \left(T'_{i,r} X_t \right) \right) = M_{X_t} \left(T_{i,r} \right), \quad i = \overline{1, m^r}$$

où $M_{X_t}(T) = \mathbb{E}(\exp\{T'X_t\})$ désigne la fonction génératrice des moments multidimensionnelle de X_t , $T_{i,r} = \frac{1}{2} \sum_{n=0}^{r-1} w_{k_{i,n}}$, et w_j étant le m -vecteur canonique $w_j = (\mathbf{0}_{1 \times (j-1)}, 1, \mathbf{0}_{1 \times m-j})'$. Ainsi,

$$\mathbb{E} \left(\boldsymbol{\varepsilon}_t^{\otimes r} \right) = \text{diag} \left\{ \left(M_{X_t} \left(T_{1,r} \right), \dots, M_{X_t} \left(T_{m^r,r} \right) \right) \right\} \mathbb{E} \left[\boldsymbol{\eta}_t^{\otimes r} \right].$$

Sous l'hypothèse de normalité du vecteur e_t , nous avons

$$\mathbb{E} \left(\boldsymbol{\varepsilon}_t^{\otimes r} \right) = \text{diag} \left\{ \left(T'_{1,r} \mu_X^{(t)} + \frac{1}{2} T'_{1,r} \Sigma_X^{(t)} T_{1,r}, \dots, T'_{m^r,r} \mu_X^{(t)} + \frac{1}{2} T'_{m^r,r} \Sigma_X^{(t)} T_{m^r,r} \right) \right\} \mathbb{E} \left[\boldsymbol{\eta}_t^{\otimes r} \right].$$

Exemple Illustratif pour le $k_{i,0}$

Soit $X = (X_1, X_2, X_3)'$ donc $m = 3$. Pour $r = 2$, on a $m^r = 9$, et

$$X^{\otimes 2} = \begin{pmatrix} X_1 & X_1 \\ X_1 & X_2 \\ X_1 & X_3 \\ X_2 & X_1 \\ X_2 & X_2 \\ X_2 & X_3 \\ X_3 & X_1 \\ X_3 & X_2 \\ X_3 & X_3 \end{pmatrix} \begin{matrix} \leftarrow (2, 0, 0) = T_{1,2} \\ \leftarrow (1, 1, 0) = T_{2,2} \\ \leftarrow (1, 0, 1) = T_{3,2} \\ \leftarrow (1, 1, 0) = T_{4,2} \\ \leftarrow (0, 2, 0) = T_{5,2} \\ \leftarrow (0, 1, 1) = T_{6,2} \\ \leftarrow (1, 0, 1) = T_{7,2} \\ \leftarrow (0, 1, 1) = T_{8,2} \\ \leftarrow (0, 0, 2) = T_{9,2} \end{matrix}$$

Pour $T_{1,2}$, on a $i = 1 = k_{i,0} + (k_{i,1} - 1)m + \dots + (k_{i,r-1} - 1)m^{r-1}$, alors $k_{1,0} = k_{1,1} = 1$, on obtient

$$T_{1,2} = \sum_{n=0}^{2-1} w_{k_{1,n}} = w_{k_{1,0}} + w_{k_{1,1}} = (1, 0, 0) + (1, 0, 0).$$

Pour $T_{7,2}$ on a ici $i = 7 = k_{7,0} + (k_{7,1} - 1) \times 3 \Rightarrow k_{7,0} = 1$ et $(k_{7,1} - 1) = 2 \Rightarrow k_{7,1} = 3$, on obtient

$$T_{7,2} = \sum_{n=0}^{2-1} w_{k_{7,n}} = w_{k_{7,0}} + w_{k_{7,1}} = (1, 0, 0) + (0, 0, 1).$$

Maintenant pour $T_{3,2}$, on a $i = 3 = k_{i,0} + (k_{i,1} - 1)m + \dots + (k_{i,r-1} - 1)m^{r-1}$, c-à-d $i = 3 = k_{i,0} + (k_{i,1} - 1) \times 3$, et puisque i dans ce cas est un multiplicateur de m alors $k_{3,0} = 3$ et $(k_{3,1} - 1) = 0$ ce qui implique que $k_{3,1} = 1$, et ce qui nous donne

$$T_{3,2} = \sum_{n=0}^{2-1} w_{k_{3,n}} = w_{k_{3,0}} + w_{k_{3,1}} = (0, 0, 1) + (1, 0, 0) = (1, 0, 1).$$

et c'est la même chose pour $T_{9,2}$, on a $i = 9 = 3 + 2 \times 3^1$ donc $k_{9,0} = 3$ et $(k_{9,1} - 1) = 0$ on obtient $k_{9,1} = 3$, ce qui nous donne

$$T_{9,2} = \sum_{n=0}^{2-1} w_{k_{9,n}} = w_{k_{9,0}} + w_{k_{9,1}} = (0, 0, 1) + (0, 0, 1) = (0, 0, 2). \blacksquare$$

Étant donné les expressions des moments d'ordre 3 et 4 ($r = 3$ et $r = 4$) du processus *PVAR-SV*, nous pouvons calculer les coefficients d'asymétrie et d'aplatissement et donner leurs expressions matricielles explicites. Ces expressions sont des fonctions des paramètres du modèle. Nous remarquons qu'en finance, la non-normalité est typiquement analysée à travers ces deux mesures définies comme suit

$$sk_{\boldsymbol{\varepsilon}}^{(t)} := \mathbb{E} \left(\left[\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{-1/2} \boldsymbol{\varepsilon}_t \right]^{\otimes 3} \right)' \mathbb{E} \left(\left[\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{-1/2} \boldsymbol{\varepsilon}_t \right]^{\otimes 3} \right),$$

et

$$\kappa_{\boldsymbol{\varepsilon}}^{(t)} := \text{vec}(\mathbf{I}_{m^2})' \mathbb{E} \left(\left[\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{-1/2} \boldsymbol{\varepsilon}_t \right]^{\otimes 4} \right),$$

où $\Lambda_{\boldsymbol{\varepsilon}}^{(t)} = \left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{1/2}$ est la racine carrée symétrique de la matrice de variance-covariance de $\boldsymbol{\varepsilon}_t$ (voir par exemple Kollo, 2008). Dans le cas d'un modèle *PVAR-SV*, nous avons

$$\begin{aligned} \mathbb{E} \left(\left[\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{-1/2} \boldsymbol{\varepsilon}_t \right]^{\otimes 3} \right) &= \left(\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{-1/2} \right)^{\otimes 3} \\ &\times \text{diag} \left\{ \left(T'_{1,3} \mu_X^{(t)} + \frac{1}{2} T'_{1,3} \Sigma_X^{(t)} T_{1,3}, \dots, T'_{m^3,3} \mu_X^{(t)} + \frac{1}{2} T'_{m^3,3} \Sigma_X^{(t)} T_{m^3,3} \right) \right\} \mathbb{E} [\boldsymbol{\eta}_t^{\otimes 3}] \\ &= \mathbf{0}_{m^3 \times 1}, \end{aligned}$$

puisque $\mathbb{E} [\boldsymbol{\eta}_t^{\otimes 3}] = \mathbf{0}_{m^3 \times 1}$, et

$$\begin{aligned} \mathbb{E} \left(\left[\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{-1/2} \boldsymbol{\varepsilon}_t \right]^{\otimes 4} \right) &= \left[\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(t)} \right)^{-1/2} \right]^{\otimes 4} \\ &\times \text{diag} \left\{ \left(T'_{1,4} \mu_X^{(t)} + \frac{1}{2} T'_{1,4} \Sigma_X^{(t)} T_{1,4}, \dots, T'_{m^4,4} \mu_X^{(t)} + \frac{1}{2} T'_{m^4,4} \Sigma_X^{(t)} T_{m^4,4} \right) \right\} \mathbb{E} [\boldsymbol{\eta}_t^{\otimes 4}]. \end{aligned}$$

Notons que sous l'hypothèse de normalité de $\boldsymbol{\eta}_t$, la forme explicite de $\mathbb{E} [\boldsymbol{\eta}_t^{\otimes r}]$ peut être obtenue à partir du résultat établi par Kollo et von Rosen (2006, Corollaire 2.2.7.4). Ainsi, les moments d'ordre impair de $\boldsymbol{\eta}_t$ sont nuls et les moments d'ordre pair sont donnés par les formules suivantes

$$\begin{aligned} \mathbb{E} (\boldsymbol{\eta}_t^{\otimes 2}) &= \text{vec} (\Sigma_{\boldsymbol{\eta}}^{(t)}), \\ \mathbb{E} (\boldsymbol{\eta}_t^{\otimes 4}) &= (\mathbf{I}_{m^4} + \mathbf{I}_m \otimes K_{m,m} \otimes \mathbf{I}_m + \mathbf{I}_m \otimes K_{m^2,m}) [\text{vec} (\Sigma_{\boldsymbol{\eta}}^{(t)})]^{\otimes 2}, \end{aligned}$$

et plus généralement

$$\mathbb{E} (\boldsymbol{\eta}_t^{\otimes r}) = \sum_{i=2}^r (\mathbf{I}_m \otimes K_{m^{i-2},m} \otimes \mathbf{I}_{m^{r-i}}) [\text{vec} (\Sigma_{\boldsymbol{\eta}}^{(t)}) \otimes \mathbb{E} (\boldsymbol{\eta}_t^{\otimes r-2})], \text{ for } r = 2, 4, 6, \dots$$

où $K_{p,q}$ est la $(pq \times pq)$ -matrice de commutation.

Proposition 3.1 (Boussaha et al., 2017) *Pour une solution p.s.s. $\{\boldsymbol{\varepsilon}_t; t \in \mathbb{Z}\}$ du modèle PAR-SV défini par (3.1) et sous l'hypothèse de normalité de $\boldsymbol{\eta}_t$ et \mathbf{e}_t , nous avons pour tout entier positif r*

$$\begin{aligned} \mu_{\boldsymbol{\varepsilon}^r}^{(s)} &:= \mathbb{E}(\boldsymbol{\varepsilon}_{s+\tau S}^{\otimes r}) \\ &= \text{diag} \left\{ \left(T'_{1,r} \mu_X^{(s)} + \frac{1}{2} T'_{1,r} \Sigma_X^{(s)} T_{1,r}, \dots, T'_{m^r,r} \mu_X^{(s)} + \frac{1}{2} T'_{m^r,r} \Sigma_X^{(s)} T_{m^r,r} \right) \right\} \mathbb{E}[\boldsymbol{\eta}_{s+\tau S}^{\otimes r}]. \end{aligned}$$

Par conséquent, la variance, le coefficient de skewness et de kurtosis de la distribution de $\boldsymbol{\varepsilon}_{s+\tau S}$ sont donnés par

$$\begin{aligned} \text{vec} \left(\Sigma_{\boldsymbol{\varepsilon},0}^{(s)} \right) &:= \mathbb{E}(\boldsymbol{\varepsilon}_{s+\tau S}^{\otimes 2}) \\ &= \text{diag} \left\{ \left(T'_{1,2} \mu_X^{(s)} + \frac{1}{2} T'_{1,2} \Sigma_X^{(s)} T_{1,2}, \dots, T'_{m^2,2} \mu_X^{(s)} + \frac{1}{2} T'_{m^2,2} \Sigma_X^{(s)} T_{m^2,2} \right) \right\} \text{vec}(\Sigma_{\boldsymbol{\eta}}^{(s)}), \\ \text{sk}_{\boldsymbol{\varepsilon}}^{(s)} &= \mathbf{0}_{m^3 \times 1}, \end{aligned}$$

et

$$\begin{aligned} \kappa_{\boldsymbol{\varepsilon}}^{(s)} &= \text{vec}(\mathbf{I}_{m^2})' \left[\left(\Sigma_{\boldsymbol{\varepsilon},0}^{(s)} \right)^{-1/2} \right]^{\otimes 4} \\ &\quad \times \text{diag} \left\{ \left(T'_{1,4} \mu_X^{(s)} + \frac{1}{2} T'_{1,4} \Sigma_X^{(s)} T_{1,4}, \dots, T'_{m^4,4} \mu_X^{(s)} + \frac{1}{2} T'_{m^4,4} \Sigma_X^{(s)} T_{m^4,4} \right) \right\} \\ &\quad \times (\mathbf{I}_{m^4} + \mathbf{I}_m \otimes K_{m,m} \otimes \mathbf{I}_m + \mathbf{I}_m \otimes K_{m^2,m}) [\text{vec}(\Sigma_{\boldsymbol{\eta}}^{(s)})]^{\otimes 2}. \end{aligned}$$

3.3.2 Fonction d'autocorrélation du modèle PVAR-SV

Partant de l'hypothèse développée au paragraphe précédent, à savoir que les variables $\boldsymbol{\eta}_t$ et \mathbf{e}_t sont des variables aléatoires normales, la structure de dépendance du processus solution de (3.1) peut être analysée. Pour ce faire, nous envisageons d'examiner, tout d'abord, sa fonction d'autocovariance.

Rappelons que la fonction d'autocovariance périodique d'un processus m -varié $\{\boldsymbol{\varepsilon}_t; t \in \mathbb{Z}\}$ périodiquement stationnaire est définie par

$$\Sigma_{\boldsymbol{\varepsilon},h}^{(s)} := \text{cov}(\boldsymbol{\varepsilon}_{s+\tau S}, \boldsymbol{\varepsilon}_{s+\tau S-h}) = \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}'_{t-h}), \quad h = 0, 1, 2, \dots$$

La proposition suivante caractérise cette fonction d'autocovariance périodique dans le cas d'un processus PVAR-SV.

Proposition 3.2 (Boussaha et al., 2017) *Soit $\{\boldsymbol{\varepsilon}_t; t \in \mathbb{Z}\}$ une solution p.s.s. de (3.1). Alors $\{\boldsymbol{\varepsilon}_t; t \in \mathbb{Z}\}$ est un processus bruit blanc faible périodique d'une variance périodique donnée, pour $s = \overline{1, S}$, par*

$$\text{vec} \left(\Sigma_{\boldsymbol{\varepsilon},0}^{(s)} \right) = \text{diag} \left\{ \left(T'_{1,2} \mu_X^{(s)} + \frac{1}{2} T'_{1,2} \Sigma_X^{(s)} T_{1,2}, \dots, T'_{m^2,2} \mu_X^{(s)} + \frac{1}{2} T'_{m^2,2} \Sigma_X^{(s)} T_{m^2,2} \right) \right\} \text{vec}(\Sigma_{\boldsymbol{\eta}}^{(s)}).$$

Preuve. Pour $h > 0$, nous avons

$$\begin{aligned}
\text{vec} \left(\Sigma_{\varepsilon,0}^{(s)} \right) &= \mathbb{E} [\boldsymbol{\varepsilon}_t \otimes \boldsymbol{\varepsilon}_{t-h}] \\
&= \mathbb{E} \left[\left(\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \boldsymbol{\eta}_t \right) \otimes \left(\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \boldsymbol{\eta}_{t-h} \right) \right] \\
&= \mathbb{E} \left[\text{diag} \left\{ \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \otimes \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \right\} (\boldsymbol{\eta}_t \otimes \boldsymbol{\eta}_{t-h}) \right] \\
&= \text{diag} \left(\mathbb{E} \left[\exp \left\{ \frac{1}{2} X_t \right\} \otimes \exp \left\{ \frac{1}{2} X_{t-h} \right\} \right] \right) \mathbb{E} (\boldsymbol{\eta}_t \otimes \boldsymbol{\eta}_{t-h}) \\
&= \text{diag} \left(\mathbb{E} \left[Y_t^{(1)} \otimes Y_{t-h}^{(1)} \right] \right) (\mathbb{E} (\boldsymbol{\eta}_t) \otimes \mathbb{E} (\boldsymbol{\eta}_{t-h})) \\
&= \text{diag} \left(\mathbb{E} \left[\exp \left\{ \frac{X_{t,1} + X_{t-h,1}}{2}, \dots, \frac{X_{t,1} + X_{t-h,m}}{2}, \dots, \frac{X_{t,m} + X_{t-h,1}}{2}, \dots, \frac{X_{t,m} + X_{t-h,m}}{2} \right\} \right] \right) \\
&\times (\mathbb{E} (\boldsymbol{\eta}_t) \otimes \mathbb{E} (\boldsymbol{\eta}_{t-h})) \\
&= \mathbf{0}_{m^2 \times 1},
\end{aligned}$$

car $(\boldsymbol{\eta}_t)$ est une séquence de variables aléatoires *i.i.d* de moyenne nulle et $(X_{t,i} + X_{t-h,j})/2$, pour tout $i, j = \overline{1, m}$, est une variable gaussienne; par conséquent $\mathbb{E} \left[\exp \left\{ \frac{X_{t,i} + X_{t-h,j}}{2} \right\} \right] < \infty, \forall i, j = \overline{1, m}$. Ce qui fait de $\{\boldsymbol{\varepsilon}_t; t \in \mathbb{Z}\}$ un bruit blanc périodique, i.e. il n'y a pas de structure de dépendance linéaire entre les termes successifs de $\{\boldsymbol{\varepsilon}_t; t \in \mathbb{Z}\}$. Notons qu'il est possible que d'autres formes de dépendance non linéaires existent. Comme dans le cas univarié, nous allons étudier la structure de dépendance du processus des puissances $\{\varepsilon_t^r; t \in \mathbb{Z}, r \in \mathbb{N}^*\}$. A partir de (3.1), nous avons

$$\begin{aligned}
\mathbb{E} [\boldsymbol{\varepsilon}_t^{\otimes 2r} \otimes \boldsymbol{\varepsilon}_{t-h}^{\otimes 2r}] &= \mathbb{E} \left[\left(\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \boldsymbol{\eta}_t \right)^{\otimes 2r} \otimes \left(\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \boldsymbol{\eta}_{t-h} \right)^{\otimes 2r} \right] \\
&= \mathbb{E} \left[\left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \right]^{\otimes 2r} \boldsymbol{\eta}_t^{\otimes 2r} \right) \otimes \left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \right]^{\otimes 2r} \boldsymbol{\eta}_{t-h}^{\otimes 2r} \right) \right] \\
&= \mathbb{E} \left[\left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \right]^{\otimes 2r} \otimes \left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \right]^{\otimes 2r} \right) (\boldsymbol{\eta}_t^{\otimes 2r} \otimes \boldsymbol{\eta}_{t-h}^{\otimes 2r}) \right] \\
&= \text{diag} \left(\mathbb{E} \left[\left(\exp \left\{ \frac{1}{2} X_t \right\} \right)^{\otimes 2r} \otimes \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right)^{\otimes 2r} \right] \right) \mathbb{E} (\boldsymbol{\eta}_t^{\otimes 2r} \otimes \boldsymbol{\eta}_{t-h}^{\otimes 2r}) \\
&= \text{diag} \left(\mathbb{E} \left[Y_t^{(2r)} \otimes Y_{t-h}^{(2r)} \right] \right) \mathbb{E} (\boldsymbol{\eta}_t^{\otimes 2r} \otimes \boldsymbol{\eta}_{t-h}^{\otimes 2r}) \\
&\neq \mathbf{0}_{m^{4r} \times 1},
\end{aligned}$$

et puisque le premier élément de chacun des deux vecteurs $Y_t^{(2r)} \otimes Y_{t-h}^{(2r)}$ et $\boldsymbol{\eta}_t^{\otimes 2r} \otimes \boldsymbol{\eta}_{t-h}^{\otimes 2r}$ est donné par

$$\left(Y_t^{(2r)} \otimes Y_{t-h}^{(2r)} \right)_1 = \exp \{ r (X_{t,1} + X_{t-h,1}) \} \text{ et } (\boldsymbol{\eta}_t^{\otimes 2r} \otimes \boldsymbol{\eta}_{t-h}^{\otimes 2r})_1 = \boldsymbol{\eta}_{t,1}^{2r} \boldsymbol{\eta}_{t-h,1}^{2r},$$

et que le premier élément du vecteur $\mathbb{E} [\boldsymbol{\varepsilon}_t^{\otimes 2r} \otimes \boldsymbol{\varepsilon}_{t-h}^{\otimes 2r}]$ satisfait

$$\begin{aligned}
(\mathbb{E} [\boldsymbol{\varepsilon}_t^{\otimes 2r} \otimes \boldsymbol{\varepsilon}_{t-h}^{\otimes 2r}])_1 &= \left(\mathbb{E} \left[Y_t^{(2r)} \otimes Y_{t-h}^{(2r)} \right] \right)_1 (\mathbb{E} [\boldsymbol{\eta}_t^{\otimes 2r} \otimes \boldsymbol{\eta}_{t-h}^{\otimes 2r}])_1 \\
&= \mathbb{E} (\exp \{ r (X_{t,1} + X_{t-h,1}) \}) \mathbb{E} (\boldsymbol{\eta}_{t,1}^{2r}) \mathbb{E} (\boldsymbol{\eta}_{t-h,1}^{2r})
\end{aligned}$$

Maintenant, puisque la matrice Φ_s est triangulaire inférieure, donc la première ligne contient seulement un seul élément différent de 0, c'est le $\Phi_s^{(1,1)}$. Dans ce cas il suffit juste de reprendre ce qui a été

fait dans le deuxième chapitre, plus précisément, on prend $2r$ au lieu de r , et à la place de β_s on prend $\Phi_s^{(1,1)}$, et à la place de $x_{s+\tau S}$, on prend la première composante de $X_{s+\tau S}$. En effet, nous avons

$$X_{s+\tau S} = \sum_{i=0}^{k-1} \left(\prod_{j=0}^{i-1} \Phi_{s-j} \right) (\alpha_{s-i} + \mathbf{e}_{s+\tau S-i}) + \left(\prod_{j=0}^{k-1} \Phi_{s-j} \right) X_{s+\tau S-k},$$

nous obtenons alors

$$X_{s+\tau S,1} = \sum_{i=0}^{k-1} \left(\prod_{j=0}^{i-1} \Phi_{s-j}^{(1,1)} \right) (\alpha_{s-i,1} + \mathbf{e}_{s+\tau S-i,1}) + \left(\prod_{j=0}^{k-1} \Phi_{s-j}^{(1,1)} \right) X_{s+\tau S-k,1},$$

ce qui donne

$$\begin{aligned} [\mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r} \otimes \boldsymbol{\varepsilon}_{t-k}^{\otimes 2r})]_1 &= \mathbb{E}(\exp\{r(X_{t,1} + X_{t-k,1})\}) \mathbb{E}(\boldsymbol{\eta}_{t,1}^{2r}) \mathbb{E}(\boldsymbol{\eta}_{t-k,1}^{2r}) \\ &= \mathbb{E} \left(\exp \left\{ \frac{r}{2} \sum_{i=0}^{k-1} \left(\prod_{j=0}^{i-1} \Phi_{s-j}^{(1,1)} \right) (\alpha_{s-i,1} + \mathbf{e}_{s+\tau S-i,1}) \right. \right. \\ &\quad \left. \left. + r \left(1 + \prod_{j=0}^{k-1} \Phi_{s-j}^{(1,1)} \right) X_{s+\tau S-k,1} \right\} \right) \mathbb{E}(\boldsymbol{\eta}_{t,1}^{2r}) \mathbb{E}(\boldsymbol{\eta}_{t-k,1}^{2r}) \\ &= \exp \left\{ \sum_{i=0}^{k-1} \left[\left(\prod_{j=0}^{i-1} \Phi_{s-j}^{(1,1)} \right) r \alpha_{s-i,1} + \left(\prod_{j=0}^{i-1} (\Phi_{s-j}^{(1,1)})^2 \right) \frac{r^2 \sum_{\mathbf{e}}^{(s-i)} (1,1)}{2} \right] \right. \\ &\quad \left. + \left(1 + \prod_{j=0}^{k-1} \Phi_{s-j}^{(1,1)} \right) r \mu_X^{(s-k)} (1) + \left(1 + \prod_{j=0}^{k-1} \Phi_{s-j}^{(1,1)} \right)^2 \frac{r^2 \sum_X^{(s-k)} (1,1)}{2} \right\} [\mathbb{E}(\boldsymbol{\eta}_{t,1}^{2r})]^2. \end{aligned}$$

D'autres part, on a

$$\begin{aligned} \mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\varepsilon}_{t-h}^{\otimes 2r}) &= \mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\varepsilon}_{t-h}^{\otimes 2r}) \\ &= \mathbb{E} \left[\left(\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \boldsymbol{\eta}_t \right)^{\otimes 2r} \right] \otimes \mathbb{E} \left[\left(\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \boldsymbol{\eta}_{t-h} \right)^{\otimes 2r} \right] \\ &= \mathbb{E} \left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \right]^{\otimes 2r} \boldsymbol{\eta}_t^{\otimes 2r} \right) \otimes \mathbb{E} \left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \right]^{\otimes 2r} \boldsymbol{\eta}_{t-h}^{\otimes 2r} \right) \\ &= \mathbb{E} \left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \right]^{\otimes 2r} \right) \mathbb{E}(\boldsymbol{\eta}_t^{\otimes 2r}) \otimes \mathbb{E} \left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \right]^{\otimes 2r} \right) \mathbb{E}(\boldsymbol{\eta}_{t-h}^{\otimes 2r}) \\ &= \left[\mathbb{E} \left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \right]^{\otimes 2r} \right) \otimes \mathbb{E} \left(\left[\text{diag} \left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \right]^{\otimes 2r} \right) \right] [\mathbb{E}(\boldsymbol{\eta}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\eta}_{t-h}^{\otimes 2r})] \\ &= \left[\text{diag} \mathbb{E} \left(\left[\left(\exp \left\{ \frac{1}{2} X_t \right\} \right) \right]^{\otimes 2r} \right) \otimes \text{diag} \mathbb{E} \left(\left[\left(\exp \left\{ \frac{1}{2} X_{t-h} \right\} \right) \right]^{\otimes 2r} \right) \right] [\mathbb{E}(\boldsymbol{\eta}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\eta}_{t-h}^{\otimes 2r})] \\ &= \left[\text{diag} \mathbb{E}(Y_t^{(2r)}) \otimes \text{diag} \mathbb{E}(Y_{t-h}^{(2r)}) \right] [\mathbb{E}(\boldsymbol{\eta}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\eta}_{t-h}^{\otimes 2r})], \end{aligned}$$

donc le premier élément de $\mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\varepsilon}_{t-h}^{\otimes 2r})$ est donné par

$$\begin{aligned} [\mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\varepsilon}_{t-h}^{\otimes 2r})]_1 &= \left[\text{diag} \mathbb{E}(Y_t^{(2r)}) \otimes \text{diag} \mathbb{E}(Y_{t-h}^{(2r)}) \right]_1 [\mathbb{E}(\boldsymbol{\eta}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\eta}_{t-h}^{\otimes 2r})]_1 \\ &= [\mathbb{E}(\exp\{r X_{t,1}\}) \mathbb{E}(\exp\{r X_{t-h,1}\})] \mathbb{E}(\boldsymbol{\eta}_{t,1}^{2r}) \mathbb{E}(\boldsymbol{\eta}_{t-h,1}^{2r}), \end{aligned}$$

tel que

$$\begin{aligned}\mathbb{E}(\exp\{rX_{t,1}\}) &= \mathbb{E}\left(\exp\left\{r\sum_{i=0}^{k-1}\left(\prod_{j=0}^{i-1}\Phi_{s-j}^{(1,1)}\right)(\alpha_{s-i,1} + \mathbf{e}_{s+\tau S-i,1}) + r\left(\prod_{j=0}^{k-1}\Phi_{s-j}^{(1,1)}\right)X_{s+\tau S-k,1}\right\}\right) \\ &= \exp\left\{\sum_{i=0}^{k-1}\left[r\left(\prod_{j=0}^{i-1}\Phi_{s-j}^{(1,1)}\right)\alpha_{s-i,1} + \frac{r^2}{4}\left(\prod_{j=0}^{i-1}\left(\Phi_{s-j}^{(1,1)}\right)^2\right)\Sigma_{\mathbf{e}}^{(s-i)}(1,1)\right]\right. \\ &\quad \left.+ r\left(\prod_{j=0}^{k-1}\Phi_{s-j}^{(1,1)}\right)\mu_X^{(s-k)}(1) + \frac{r^2}{4}\left(\prod_{j=0}^{k-1}\left(\Phi_{s-j}^{(1,1)}\right)^2\right)\Sigma_X^{(s-k)}(1,1)\right\},\end{aligned}$$

et

$$\mathbb{E}(\exp\{rX_{s+\tau S-k,1}\}) = \exp\left\{r\mu_X^{(s-k)}(1) + \frac{r^2}{4}\Sigma_X^{(s-k)}(1,1)\right\}\mathbb{E}(\boldsymbol{\eta}_{t-k,1}^{2r}),$$

d'où

$$\begin{aligned}[\mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r} \otimes \boldsymbol{\varepsilon}_{t-k}^{\otimes 2r})]_1 &= \exp\left\{\sum_{i=0}^{k-1}\left[\left(\prod_{j=0}^{i-1}\Phi_{s-j}^{(1,1)}\right)r\alpha_{s-i,1} + \left(\prod_{j=0}^{i-1}\left(\Phi_{s-j}^{(1,1)}\right)^2\right)\frac{r^2\Sigma_{\mathbf{e}}^{(s-i)}(1,1)}{2}\right]\right. \\ &\quad \left.+ \left(1 + \prod_{j=0}^{k-1}\Phi_{s-j}^{(1,1)}\right)r\mu_X^{(s-k)}(1) + \left(1 + \prod_{j=0}^{k-1}\left(\Phi_{s-j}^{(1,1)}\right)^2\right)\frac{r^2\Sigma_X^{(s-k)}(1,1)}{2}\right\}[\mathbb{E}(\boldsymbol{\eta}_{t,1}^{2r})]^2.\end{aligned}$$

On obtient finalement,

$$\begin{aligned}&[\mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r} \otimes \boldsymbol{\varepsilon}_{t-h}^{\otimes 2r})]_1 - [\mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r}) \otimes \mathbb{E}(\boldsymbol{\varepsilon}_t^{\otimes 2r})]_1 \\ &= [\mathbb{E}(\boldsymbol{\eta}_{t,1}^{2r})]^2 \exp\left\{\sum_{i=0}^{k-1}\left[\left(\prod_{j=0}^{i-1}\Phi_{s-j}^{(1,1)}\right)r\alpha_{s-i} + \left(\prod_{j=0}^{i-1}\left(\Phi_{s-j}^{(1,1)}\right)^2\right)\frac{r^2\Sigma_{\mathbf{e}}^{(s-i)}(1,1)}{2}\right]\right. \\ &\quad \left.+ \left(1 + \prod_{j=0}^{k-1}\Phi_{s-j}^{(1,1)}\right)r\mu_X^{(s-k)}(1) + \left(1 + \prod_{j=0}^{k-1}\left(\Phi_{s-j}^{(1,1)}\right)^2\right)\frac{r^2\Sigma_X^{(s-k)}(1,1)}{2}\right\} \\ &\quad \times \left[\exp\left\{\left(\prod_{j=0}^{k-1}\Phi_{s-j}^{(1,1)}\right)\frac{r^2\Sigma_X^{(s-k)}(1,1)}{2}\right\} - \exp\{1\}\right] \neq 0.\end{aligned}$$

■

3.4 Estimation des paramètres d'un modèle *PVAR-SV*

Dans cette section, nous allons adopter les mêmes méthodologies d'estimation proposées dans le cas *PAR-SV*. Nous décrivons plus particulièrement la seconde méthode qui s'avère plus utile dans la pratique.

3.4.1 Méthode du quasi-maximum de vraisemblance basée sur le filtre de Kalman périodique

Afin d'estimer les paramètres du modèle *SV multivarié non-périodique*, Harvey et al. (1994) ont utilisé la méthode du quasi-maximum de vraisemblance (*QML*) qui consiste en la linéarisation de l'équation d'observation après avoir défini la variable observée comme un logarithme du vecteur

$\boldsymbol{\varepsilon}_t \odot \boldsymbol{\varepsilon}_t$, où l'opérateur \odot désigne le produit d'Hadamard (produit élément par élément). Dans notre cas périodique, le modèle (3.1) peut être linéarisé en prenant la même transformation $\ln(\boldsymbol{\varepsilon}_{s+\tau S} \odot \boldsymbol{\varepsilon}_{s+\tau S})$. Ainsi, le modèle obtenu admet la représentation espace d'états suivante

$$\begin{cases} Z_{s+\tau S} = X_{s+\tau S} + \mathbf{d}_s + \mathbf{u}_{s+\tau S}, \\ X_{s+\tau S} = \boldsymbol{\alpha}_s + \Phi_s X_{s+\tau S-1} + \mathbf{e}_{s+\tau S}, \end{cases} \quad \tau \in \mathbb{Z}, \quad 1 \leq s \leq S, \quad (3.5)$$

où $Z_{s+\tau S} = (Z_{s+\tau S,1}, Z_{s+\tau S,2}, \dots, Z_{s+\tau S,m})'$, $Z_{s+\tau S,i} = \ln(\boldsymbol{\varepsilon}_{s+\tau S,i}^2)$, $\mathbf{d}_s = \mathbb{E}[\ln(\boldsymbol{\eta}_{s+\tau S} \odot \boldsymbol{\eta}_{s+\tau S})]$ et $\mathbf{u}_{s+\tau S} = \ln(\boldsymbol{\eta}_{s+\tau S} \odot \boldsymbol{\eta}_{s+\tau S}) - \mathbb{E}[\ln(\boldsymbol{\eta}_{s+\tau S} \odot \boldsymbol{\eta}_{s+\tau S})]$. Il est clair que même si le processus aléatoire $\boldsymbol{\eta}_{s+\tau S}$ est gaussien, les erreurs de l'équation d'observation, $\mathbf{u}_{s+\tau S}$, dans (3.5) ne le sont pas. La moyenne \mathbf{d}_s et l'élément (i, j) de la matrice de variance-covariance $\Sigma_u^{(s)}$ du vecteur aléatoire $\mathbf{u}_{s+\tau S}$ peuvent être donnés par

$$\mathbf{d}_s = (-1.2749 + \Sigma_\eta^{(s)}(1, 1), \dots, -1.2749 + \Sigma_\eta^{(s)}(m, m))',$$

et

$$\Sigma_u^{(s)}(i, j) = \begin{cases} \sum_{n=1}^{\infty} \frac{(n-1)!}{n(\prod_{k=1}^n (\frac{1}{2} + k - 1))} \left(\frac{\Sigma_\eta^{(s)}(i, j)}{\sqrt{\Sigma_\eta^{(s)}(i, i)} \sqrt{\Sigma_\eta^{(s)}(j, j)}} \right)^{2n}, & \text{si } i \neq j, \quad i, j = \overline{1, m}, \\ \frac{\pi^2}{2}, & \text{sinon,} \end{cases}$$

(voir pour le cas non périodique Harvey et al, 1994). La représentation (3.5) rend le problème d'estimation par la méthode *QML* plus évident (voir Harvey et al. (1994) pour le cas multivarié non périodique et Boussaha et Hamdi (2017) pour le cas univarié périodique). Cependant, dans le cas multivarié, la transformation du modèle entraîne une perte d'information. Plus précisément, Harvey et al. (1994) ont noté qu'on pourrait faire qu'une estimation des valeurs absolues des paramètres inconnus $\Sigma_\eta(i, j)$, $i, j = \overline{1, m}$. Il s'agit des corrélations croisées entre les différentes composantes $\eta_{t,i}$, mais leurs signes ne peuvent pas être estimés. Afin de pallier ce problème, Harvey et al. (1994) ont suggéré d'utiliser les signes des observations non transformées pour obtenir les signes des coefficients de corrélation. Ils ont proposé d'estimer le signe de $\Sigma_\eta(i, j)$ comme positif dans le cas où plus de la moitié des paires $\boldsymbol{\varepsilon}_{t,i} \boldsymbol{\varepsilon}_{t,j}$ sont positives. Nous adoptons la même idée dans notre cas périodique, i.e. le signe de $\Sigma_\eta^{(s)}(i, j)$ est estimé comme positif si plus de la moitié des paires $\boldsymbol{\varepsilon}_{s+\tau S,i} \boldsymbol{\varepsilon}_{s+\tau S,j}$ sont positives.

Pour discuter la méthode *QML* basée sur le filtre de Kalman périodique, procédons comme dans le cas univarié. Soit $\widehat{Z}_{s+\tau S|s+\tau S-1}$ le meilleur prédicteur linéaire (au sens des moindres carrés) de $Z_{s+\tau S}$ basé sur $Z_1, Z_2, \dots, Z_{s+\tau S-1}$, et $\widehat{\mathbf{u}}_{s+\tau S} = Z_{s+\tau S} - \widehat{Z}_{s+\tau S|s+\tau S-1}$ est l'innovation empirique à l'instant $s + \tau S$, dont l'erreur moyenne quadratique est $\Omega_{s+\tau S} = \mathbb{E}(\widehat{\mathbf{u}}_{s+\tau S} \widehat{\mathbf{u}}_{s+\tau S}')$. L'innovation est donnée aussi par $\widehat{\mathbf{u}}_{s+\tau S} = Z_{s+\tau S} - \widehat{X}_{s+\tau S|s+\tau S-1} - \mathbf{d}_s$, où $\widehat{X}_{s+\tau S|s+\tau S-1}$ est le meilleur prédicteur linéaire de $X_{s+\tau S}$ basé sur $Z_1, Z_2, \dots, Z_{s+\tau S-1}$, dont la matrice de covariance de l'erreur de prédiction est

$$P_{s+\tau S|s+\tau S-1} = \mathbb{E} \left[\left(X_{s+\tau S} - \widehat{X}_{s+\tau S|s+\tau S-1} \right) \left(X_{s+\tau S} - \widehat{X}_{s+\tau S|s+\tau S-1} \right)' \right].$$

Pour une réalisation donnée $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_n)$ du modèle (3.1), la fonction quasi-vraisemblance du vecteur des paramètres $\boldsymbol{\theta} = (\boldsymbol{\theta}'_1, \dots, \boldsymbol{\theta}'_S)'$, où $\boldsymbol{\theta}_s = \left(\boldsymbol{\alpha}_s, (\text{vech}(\Phi_s))', (\text{vech}(\Sigma_{\mathbf{e}}^{(s)}))', (\text{vech}(\Sigma_{\boldsymbol{\eta}}^{(s)}))' \right)'$, peut être écrite sous la forme d'innovation suivante

$$L(\boldsymbol{\theta}; Z) = (2\pi)^{-\frac{nm}{2}} \prod_{s=1}^S \prod_{\tau=0}^{\tau_1-1} (\det \Omega_{s+\tau S})^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \widehat{\mathbf{u}}'_{s+\tau S} \Omega_{s+\tau S}^{-1} \widehat{\mathbf{u}}_{s+\tau S} \right\}, \quad (3.6)$$

dans laquelle nous avons besoin d'évaluer $\Omega_{s+\tau S}$ et $\widehat{\mathbf{u}}_{s+\tau S}$, pour $s = \overline{1, S}$ et $\tau = \overline{0, \tau_1 - 1}$. Comme dans le cas univarié et sans perdre de généralité, nous supposons que le nombre d'observations est un multiple de la période, i.e. $n = \tau_1 S$.

La fonction quasi-vraisemblance (3.6) peut être évaluée en utilisant l'algorithme du filtre de Kalman périodique ou l'algorithme de Chandrasekhar périodique (Aknouche et Hamdi, 2007). D'une façon simple, et à partir de la représentation espace d'états (3.5), le filtre périodique de Kalman, associé à notre modèle, est donné par les récursions suivantes

$$\begin{cases} \widehat{\mathbf{u}}_{s+\tau S} &= Z_{s+\tau S} - \widehat{X}_{s+\tau S|s+\tau S-1} - \mathbf{d}_s, \\ \Omega_{s+\tau S} &= P_{s+\tau S|s+\tau S-1} + \Sigma_{\mathbf{u}}^{(s)}, \\ K_{s+\tau S} &= \Phi_{s+1} P_{s+\tau S|s+\tau S-1} \Omega_{s+\tau S}^{-1}, \\ \widehat{X}_{s+\tau S+1|s+\tau S} &= \boldsymbol{\alpha}_{s+1} + \Phi_{s+1} \widehat{X}_{s+\tau S|s+\tau S-1} + K_{s+\tau S} \widehat{\mathbf{u}}_{s+\tau S}, \\ P_{s+\tau S+1|s+\tau S} &= \Phi_{s+1} \left(P_{s+\tau S|s+\tau S-1} - P_{s+\tau S|s+\tau S-1} \Omega_{s+\tau S}^{-1} P_{s+\tau S|s+\tau S-1} \right) \Phi'_{s+1} + \Sigma_{\mathbf{e}}^{(s+1)}, \end{cases}$$

avec les valeurs de démarrage

$$\widehat{X}_{1|0} = \left(\mathbf{I}_m - \prod_{v=0}^{S-1} \Phi_{s-v} \right)^{-1} \sum_{v=0}^{S-1} \left(\prod_{i=0}^{v-1} \Phi_{s-i} \right) \boldsymbol{\alpha}_{s-v},$$

et

$$\begin{aligned} \text{vec}(P_{1|0}) &= \left(\mathbf{I}_{m^2} - \left(\prod_{v=0}^{S-1} \Phi_{1-v} \right)^{\otimes 2} \right)^{-1} \\ &\times \sum_{v=0}^{S-1} \left(\prod_{i=0}^{v-1} \Phi_{1-i} \right)^{\otimes 2} \left[\left(\boldsymbol{\alpha}_{1-v} - \mu_X^{(1-v)} + \Phi_{1-v} \mu_X^{(S-v)} \right)^{\otimes 2} + \text{vec}(\Sigma_{\mathbf{e}}^{(1-v)}) \right]. \end{aligned}$$

En maximisant la fonction quasi-vraisemblance (3.6) par rapport au vecteur $\boldsymbol{\theta}$, nous obtenons $\widehat{\boldsymbol{\theta}}$ l'estimateur *QML* de $\boldsymbol{\theta}$. Notons qu'il n'existe pas d'expression explicite de $\widehat{\boldsymbol{\theta}}$. Pour cette raison, nous faisons appel à des méthodes d'optimisation numériques. Notons enfin que ces résultats nous ont conduits à définir une procédure en quatre étapes pour estimer les paramètres du modèle *PVAR-SV* par la méthode *QML* basée sur le filtre de Kalman périodique. Ces quatre étapes peuvent être récapitulées dans l'algorithme suivant.

Algorithme 3.1 (Méthode QML basée sur le filtre de Kalman périodique)

1. Utiliser un estimateur préliminaire $\widehat{\theta}^{(0)} = (\widehat{\theta}_1^{(0)'}, \dots, \widehat{\theta}_S^{(0)'})'$ du vecteur θ , où

$$\widehat{\theta}_s^{(0)} = \left(\widehat{\alpha}_s^{(0)}, \left(\text{vech} \left(\widehat{\Phi}_s^{(0)} \right) \right)', \left(\text{vech} \left(\widehat{\Sigma}_e^{(s,0)} \right) \right)', \left(\text{vech} \left(\widehat{\Sigma}_\eta^{(s,0)} \right) \right)' \right)', \quad s = \overline{1, S}.$$

2. Exécuter le filtre de Kalman périodique pour évaluer $\widehat{\mathbf{u}}_{s+\tau S}$ et $\Omega_{s+\tau S}$, pour $s = \overline{1, S}, \tau = \overline{0, \tau_1}$.

3. Evaluer la fonction QML, $L(\theta; Z)$, à partir de (3.6).

4. Utiliser une procédure d'optimisation pour déterminer $\widehat{\theta}$ qui maximise $L(\theta; Z)$.

Pour les mêmes raisons données dans le chapitre précédent, nous adoptons aussi le filtre particulière et l'algorithme *EM* comme une deuxième méthode d'estimation d'un modèle *PVAR-SV*.

3.4.2 Méthode du maximum de vraisemblance basée sur l'algorithme *EM* et le filtre particulière

Dans ce paragraphe, nous décrivons une deuxième approche pour le problème de l'estimation, qui est une généralisation de la méthode décrite dans le chapitre précédent. Soient $\underline{Y} = (X'_0, X'_1, \dots, X'_n, \varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_n)'$ le vecteur des données complètes et $\underline{X} = (X'_0, X'_1, \dots, X'_n)$ le vecteur des log-volatilités.

Pour une réalisation donnée $Y = (\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_n)$ du modèle (3.1), la fonction de vraisemblance complète du vecteur des paramètres θ , peut être exprimée comme suit

$$\begin{aligned} \mathbf{L}(\theta; \underline{Y}) &= f(X_0) \prod_{t=1}^n f(X_t | \underline{X}_{t-1}) f(\varepsilon_t | \underline{\varepsilon}_{t-1}, \underline{X}) \\ &= f(X_0) \prod_{t=1}^n f(X_t | X_{t-1}) \prod_{t=1}^n f(\varepsilon_t | X_t) \\ &= f(X_0) \prod_{t=1}^n \frac{1}{\left((2\pi)^m \det(\Sigma_e^{(t)}) \right)^{1/2}} \exp \left(- \frac{(X_t - \alpha_t - \Phi_t X_{t-1})' \left(\Sigma_e^{(t)} \right)^{-1} (X_t - \alpha_t - \Phi_t X_{t-1})}{2} \right) \\ &\quad \times \prod_{t=1}^n \frac{\prod_{i=1}^m \exp \left(- \frac{X_i^{(t)}}{2} \right)}{\left((2\pi)^m \det(\Sigma_\eta^{(t)}) \right)^{1/2}} \\ &\quad \times \exp \left(- \frac{(\text{diag}(\exp \{ -\frac{1}{2} X_t \}) \varepsilon_t)' \left(\Sigma_\eta^{(t)} \right)^{-1} (\text{diag}(\exp \{ -\frac{1}{2} X_t \}) \varepsilon_t)}{2} \right). \end{aligned}$$

telle que $\underline{X}_{t-1} = (X'_0, X'_1, \dots, X'_{t-1})$ et $\underline{\varepsilon}_{t-1} = (\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_{t-1})$. Ainsi, la fonction log-vraisemblance

complète du vecteur des paramètres θ est donnée par

$$\begin{aligned}
-2 \log \mathbf{L}(\theta; \underline{Y}) &= C_3 + \sum_{t=1}^n \log(\det(\Sigma_{\mathbf{e}}^{(t)})) + \sum_{t=1}^n (X_t - \boldsymbol{\alpha}_t - \Phi_t X_{t-1})' (\Sigma_{\mathbf{e}}^{(t)})^{-1} (X_t - \boldsymbol{\alpha}_t - \Phi_t X_{t-1}) \\
&\quad + \sum_{t=1}^n \log(\det(\Sigma_{\boldsymbol{\eta}}^{(t)})) + \sum_{t=1}^n \sum_{i=1}^m X_{t,i} \\
&\quad + \sum_{t=1}^n \left(\text{diag} \left(\exp \left\{ -\frac{1}{2} X_t \right\} \right) \boldsymbol{\varepsilon}_t \right)' (\Sigma_{\boldsymbol{\eta}}^{(t)})^{-1} \left(\text{diag} \left(\exp \left\{ -\frac{1}{2} X_t \right\} \right) \boldsymbol{\varepsilon}_t \right),
\end{aligned}$$

où C_3 est une constante indépendante du θ . Afin d'utiliser l'algorithme *EM*, définissons la fonction \mathbf{Q}

$$\begin{aligned}
\mathbf{Q}(\theta, \widehat{\theta}^{(i)}) &= \mathbb{E} \left[-2 \log \mathbf{L}(\theta; \underline{Y}) | Y, \widehat{\theta}^{(i)} \right] \\
&= C + \sum_{t=1}^n \log(\det(\Sigma_{\mathbf{e}}^{(t)})) + \sum_{t=1}^n \log(\det(\Sigma_{\boldsymbol{\eta}}^{(t)})) + \sum_{t=1}^n \sum_{i=1}^m X_{t,i}^{(n)} \\
&\quad + \sum_{t=1}^n \text{tr} \left((\Sigma_{\mathbf{e}}^{(t)})^{-1} A_t \right) + \sum_{t=1}^n \text{tr} \left((\Sigma_{\boldsymbol{\eta}}^{(t)})^{-1} B_t \right),
\end{aligned}$$

où

$$\begin{aligned}
A_t &:= \mathbb{E} \left[(X_t - \boldsymbol{\alpha}_t - \Phi_t X_{t-1}) (X_t - \boldsymbol{\alpha}_t - \Phi_t X_{t-1})' | Y \right] \\
&= X_t^{(n)} \left(X_t^{(n)} \right)' + P_t^{(n)} - X_t^{(n)} \boldsymbol{\alpha}_t' - \left(X_t^{(n)} \left(X_{t-1}^{(n)} \right)' + P_{t,t-1}^{(n)} \right) \Phi_t' \\
&\quad - \boldsymbol{\alpha}_t \left(X_t^{(n)} \right)' + \boldsymbol{\alpha}_t \boldsymbol{\alpha}_t' + \boldsymbol{\alpha}_t \left(X_{t-1}^{(n)} \right)' \Phi_t' \\
&\quad - \Phi_t \left(X_{t-1}^{(n)} \left(X_t^{(n)} \right)' + P_{t,t-1}^{(n)} \right) + \Phi_t X_{t-1}^{(n)} \boldsymbol{\alpha}_t' + \Phi_t \left(X_{t-1}^{(n)} \left(X_{t-1}^{(n)} \right)' + P_{t-1}^{(n)} \right) \Phi_t',
\end{aligned}$$

$$B_t := \mathbb{E} \left[\left(\text{diag} \left(\exp \left\{ -\frac{1}{2} X_t \right\} \right) \boldsymbol{\varepsilon}_t \right) \left(\text{diag} \left(\exp \left\{ -\frac{1}{2} X_t \right\} \right) \boldsymbol{\varepsilon}_t \right)' | Y \right],$$

$$X_t^{(n)} := \mathbb{E} \left(X_t | Y, \widehat{\theta}^{(i)} \right),$$

$$P_t^{(n)} := \mathbb{E} \left(\left(X_t - X_t^{(n)} \right) \left(X_t - X_t^{(n)} \right)' | Y, \widehat{\theta}^{(i)} \right),$$

et

$$P_{t,t-1}^{(n)} := \mathbb{E} \left(\left(X_t - X_t^{(n)} \right) \left(X_{t-1} - X_{t-1}^{(n)} \right)' | Y, \widehat{\theta}^{(i)} \right).$$

Avant de passer à l'étape M , nous devons évaluer les quantités $X_t^{(n)}$, $P_t^{(n)}$, $P_{t,t-1}^{(n)}$, A_t et B_t . Dans le modèle *SV* univarié (voir Kim et Stoffer, 2008 et Boussaha et Hamdi, 2017), ces quantités ont été approximées séquentiellement à l'aide des algorithmes de filtrage et de lissage particulière. L'algorithme de l'étape de filtrage nous permet d'obtenir un échantillon de taille M distribué selon la densité de probabilité $p(X_t | \mathcal{F}_t)$, où \mathcal{F}_t désigne la σ algèbre basée sur les informations disponibles jusqu'à l'instant t . L'algorithme proposé ci-dessous vise à généraliser l'Algorithme 2.4 présenté dans le chapitre précédent au cas multivarié.

Algorithme 3.2 (Algorithme de filtrage particulière pour un modèle PVAR-SV)

1. Initialisation : pour $j = \overline{1, M}$, générer $f_0^{(j)} \sim p_0(x_0)$ avec les poids initiaux $w_0^{(j)} = 1/M$, où M est le nombre de particules.

2. Pour $s = \overline{1, S}$ et $\tau = \overline{0, \tau_1 - 1}$.

(a) Pour $j = \overline{1, M}$

i. Générer $\mathbf{e}_{s+\tau S}^{(j)} \sim \mathcal{N}(\mathbf{0}_{m \times 1}, \Sigma_{\mathbf{e}}^{(s)})$.

ii. Calculer $p_{s+\tau S}^{(j)} = \boldsymbol{\alpha}_s + \Phi_s f_{s+\tau S-1}^{(j)} + \mathbf{e}_{s+\tau S}^{(j)}$.

iii. Mettre à jour les poids : calculer

$$\begin{aligned} w_{s+\tau S}^{(j)} &= w_{s+\tau S-1}^{(j)} p\left(\mathbf{e}_{s+\tau S} \mid p_{s+\tau S}^{(j)}\right) \\ &\propto w_{s+\tau S-1}^{(j)} \exp \left\{ -\frac{\left(\text{diag}\left(\exp\left(-\frac{1}{2} p_{s+\tau S}^{(j)}\right)\right)\right) \mathbf{e}_{s+\tau S}' \left(\Sigma_{\boldsymbol{\eta}}^{(s)}\right)^{-1} \left(\text{diag}\left(\exp\left(-\frac{1}{2} p_{s+\tau S}^{(j)}\right)\right)\right) \mathbf{e}_{s+\tau S}}{2} \right\} \\ &\quad \times \left(\det\left(\Sigma_{\boldsymbol{\eta}}^{(s)}\right)\right)^{-1/2} \det\left(\text{diag}\left(\exp\left(-\frac{1}{2} p_{s+\tau S}^{(j)}\right)\right)\right). \end{aligned}$$

(b) Pour $j = \overline{1, M}$, normaliser les poids : calculer $\tilde{w}_{s+\tau S}^{(j)} = w_{s+\tau S}^{(j)} / \sum_{j=1}^M w_{s+\tau S}^{(j)}$.

(c) Calculer la mesure de dégénérescence $n^{eff} = 1 / \sum_{j=1}^M \tilde{w}_{s+\tau S}^{(j)2}$. Si $n^{eff} \leq n^T$ rééchantillonner $\left\{ \left(p_{s+\tau S}^{(j)}, \tilde{w}_{s+\tau S}^{(j)} \right), j = \overline{1, M} \right\}$ pour obtenir M particules de pondérations égales, $\left\{ \left(f_{s+\tau S}^{(j)}, 1/M \right), j = \overline{1, M} \right\}$.

3. Finalement, la séquence des M particules $\left\{ f_{s+\tau S}^{(j)}; j = \overline{1, M} \right\}$ est un échantillon aléatoire de $p(X_{s+\tau S} | \mathcal{F}_{s+\tau S})$ pour $s = \overline{1, S}$ et $\tau = \overline{0, \tau_1 - 1}$.

L'algorithme de lissage particulière suivant nous permet d'approximer des quantités $X_t^{(n)}, P_t^{(n)}, P_{t,t-1}^{(n)}$ et B_t , qui sont nécessaires dans l'étape E .

Algorithme 3.3 (Algorithme de lissage particulière pour un modèle PVAR-SV)

1. Pour $j = \overline{1, M}$, choisir $s_n^{(j)} = f_n^{(i)}$, avec les probabilités $\tilde{w}_n^{(i)}$. Poser $W_n^{(j)} = 1/M$.

2. Pour $j = \overline{1, M}$.

(a) Pour $s = S, \dots, 1$ et $\tau = \tau_1 - 1, \dots, 0$, calculer pour $i = \overline{1, M}$

$$\begin{aligned} W_{s+\tau S-1|s+\tau S}^{(i)} &= \tilde{w}_{s+\tau S-1}^{(i)} \exp \left\{ -\frac{\left(s_{s+\tau S}^{(j)} - \boldsymbol{\alpha}_s - \Phi_s f_{s+\tau S-1}^{(i)}\right)' \left(\Sigma_{\mathbf{e}}^{(s)}\right)^{-1} \left(s_{s+\tau S}^{(j)} - \boldsymbol{\alpha}_s - \Phi_s f_{s+\tau S-1}^{(i)}\right)}{2} \right\} \\ &\quad \times \left(\det\left(\Sigma_{\mathbf{e}}^{(s)}\right)\right)^{-1/2}. \end{aligned}$$

(b) Pour $i = \overline{1, M}$, normaliser les poids lissés : calculer

$$\tilde{W}_{s+\tau S-1|s+\tau S}^{(i)} = W_{s+\tau S-1|s+\tau S}^{(i)} / \sum_{j=1}^M W_{s+\tau S-1|s+\tau S}^{(i)}.$$

(c) Choisir $s_{s+\tau S-1}^{(j)} = f_{s+\tau S-1}^{(i)}$, avec les probabilités $\widetilde{W}_{s+\tau S-1|s+\tau S}^{(i)}$.

(d) Finalement, calculer pour $s = \overline{1, S}$ et $\tau = \overline{0, \tau_1 - 1}$,

$$\begin{aligned}\widehat{X}_{s+\tau S}^{(n)} &= \frac{\sum_{j=1}^M s_{s+\tau S}^{(j)}}{M}, \\ \widehat{P}_{s+\tau S}^{(n)} &= \frac{\sum_{j=1}^M \left(s_{s+\tau S}^{(j)} - \widehat{X}_{s+\tau S}^{(n)} \right) \left(s_{s+\tau S}^{(j)} - \widehat{X}_{s+\tau S}^{(n)} \right)'}{M-1}, \\ \widehat{P}_{s+\tau S, s+\tau S-1}^{(n)} &= \frac{\sum_{j=1}^M \left(s_{s+\tau S}^{(j)} - \widehat{X}_{s+\tau S}^{(n)} \right) \left(s_{s+\tau S-1}^{(j)} - \widehat{X}_{s+\tau S-1}^{(n)} \right)'}{M},\end{aligned}$$

et

$$\widehat{B}_t = \frac{\sum_{j=1}^M \left[\text{diag} \left(\exp \left\{ -\frac{1}{2} s_t^{(j)} \right\} \right) \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t' \text{diag} \left(\exp \left\{ -\frac{1}{2} s_t^{(j)} \right\} \right) \right]}{M}.$$

Une fois $X_t^{(n)}$, $P_t^{(n)}$ et $P_{t,t-1}^{(n)}$ remplacées par leurs approximations, nous passons à l'étape M , où $\widehat{\theta}^{(i+1)}$ est obtenu par

$$\theta^{(i+1)} = \arg \max_{\theta} Q \left(\theta, \widehat{\theta}^{(i)} \right).$$

Ceci peut être accompli par la résolution du système d'équations suivant

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial \boldsymbol{\alpha}_s} = 0, \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial (\text{vech}(\Phi_s))} = 0, \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial (\text{vech}(\Sigma_{\mathbf{e}}^{(s)}))} = 0, \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial (\text{vech}(\Sigma_{\eta}^{(s)}))} = 0, \end{array} \right. \quad \text{pour tout } s = \overline{1, S}.$$

Les dérivées premières de $\mathbf{Q}(\theta, \widehat{\theta}^{(i)})$ par rapport à $\boldsymbol{\alpha}_s$, $\text{vech}(\Phi_s)$, $\text{vech}(\Sigma_{\mathbf{e}}^{(s)})$ et $\text{vech}(\Sigma_{\eta}^{(s)})$, sont respectivement données, pour $s = \overline{1, S}$, par

$$\left\{ \begin{array}{l} \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial \boldsymbol{\alpha}_s} = \left(\Sigma_{\mathbf{e}}^{(s)} \right)^{-1} \sum_{\tau=0}^{\tau_1-1} \left(2\boldsymbol{\alpha}_s - 2X_{s+\tau S}^{(n)} + 2\Phi_s X_{s+\tau S-1}^{(n)} \right), \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial (\text{vech}(\Phi_s))} = \text{vech} \left(\frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial \Phi_s} \right), \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial (\text{vech}(\Sigma_{\mathbf{e}}^{(s)}))} = D'_m \times \text{vec} \left(\frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial \Sigma_{\mathbf{e}}^{(s)}} \right), \\ \frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial (\text{vech}(\Sigma_{\eta}^{(s)}))} = D'_m \times \text{vec} \left(\frac{\partial \mathbf{Q}(\theta, \widehat{\theta}^{(i)})}{\partial \Sigma_{\eta}^{(s)}} \right), \end{array} \right.$$

où

$$\begin{aligned}\frac{\partial \mathbf{Q}(\theta, \hat{\theta}^{(i)})}{\partial \Phi_s} &= 2 (\Sigma_{\mathbf{e}}^{(s)})^{-1} \sum_{\tau=0}^{\tau_1-1} \left[\alpha_s \left(X_{s+\tau S-1}^{(n)} \right)' - \left(X_{s+\tau S}^{(n)} \left(X_{s+\tau S-1}^{(n)} \right)' + P_{s+\tau S, s+\tau S-1}^{(n)} \right) \right. \\ &\quad \left. + \Phi_s \left(X_{s+\tau S-1}^{(n)} \left(X_{s+\tau S-1}^{(n)} \right)' + P_{s+\tau S-1}^{(n)} \right) \right], \\ \frac{\partial \mathbf{Q}(\theta, \hat{\theta}^{(i)})}{\partial \Sigma_{\mathbf{e}}^{(s)}} &= (\Sigma_{\mathbf{e}}^{(s)})^{-1} \sum_{\tau=0}^{\tau_1-1} \left[\mathbf{I}_m - A_{s+\tau S} (\Sigma_{\mathbf{e}}^{(s)})^{-1} \right], \\ \frac{\partial \mathbf{Q}(\theta, \hat{\theta}^{(i)})}{\partial \Sigma_{\eta}^{(s)}} &= (\Sigma_{\eta}^{(s)})^{-1} \sum_{\tau=0}^{\tau_1-1} \left[\mathbf{I}_m - B_{s+\tau S} (\Sigma_{\eta}^{(s)})^{-1} \right],\end{aligned}$$

et D_m désigne la $(m^2 \times \frac{1}{2}m(m+1))$ -matrice de duplication définie par $vec(A) = D_m \times vech(A)$, pour toute matrice A de dimension $(m \times m)$ (voir par exemple Lütkepohl, 1996, Section 9.5). Ici $vech(B)$ est l'opérateur qui permet de stocker dans une colonne les éléments situés sur et en-dessous de la diagonale de la matrice symétrique B .

Par conséquent, à la $(i+1)^{ème}$ itération, les estimateurs des paramètres de α_s , $vech(\Phi_s)$, $vech(\Sigma_{\mathbf{e}}^{(s)})$ et $vech(\Sigma_{\eta}^{(s)})$, pour $s = \overline{1, S}$, sont respectivement donnés par

$$\left\{ \begin{aligned} vech(\hat{\Sigma}_{\eta}^{(s)}) &= \frac{1}{\tau_1} \sum_{\tau=0}^{\tau_1-1} vech(\hat{B}_{s+\tau S}), \\ vech(\hat{\Phi}_s) &= \left[L_m \left\{ \left(\frac{1}{\tau_1} \left(\sum_{\tau=0}^{\tau_1-1} X_{s+\tau S-1}^{(n)} \right) \left(\sum_{\tau=0}^{\tau_1-1} X_{s+\tau S-1}^{(n)} \right)' \right. \right. \right. \\ &\quad \left. \left. - \sum_{\tau=0}^{\tau_1-1} \left(X_{s+\tau S-1}^{(n)} \left(X_{s+\tau S-1}^{(n)} \right)' + P_{s+\tau S-1}^{(n)} \right) \right) \otimes \mathbf{I}_m \right\} L_m' \right]^{-1}, \\ \hat{\alpha}_s &= \frac{1}{\tau_1} \sum_{\tau=0}^{\tau_1-1} X_{s+\tau S}^{(n)} - \frac{1}{\tau_1} \hat{\Phi}_s \sum_{\tau=0}^{\tau_1-1} X_{s+\tau S-1}^{(n)}, \\ vech(\hat{\Sigma}_{\mathbf{e}}^{(s)}) &= \frac{1}{\tau_1} \sum_{\tau=0}^{\tau_1-1} vech(\hat{A}_{s+\tau S}), \end{aligned} \right. \quad (3.7)$$

où L_m désigne la $(\frac{1}{2}m(m+1) \times m^2)$ -matrice d'élimination définie par $vech(A) = L_m \times vec(A)$, pour toute matrice A de dimension $(m \times m)$ (voir par exemple Lütkepohl, 1996, Section 9.6).

L'algorithme suivant récapitule la procédure d'estimation par la méthode du maximum de vraisemblance, des paramètres d'un modèle $PVAR-SV$, en se basant sur l'algorithme EM et le filtre particulière.

Algorithme 3.4 (Méthode ML basée sur l'algorithme EM et le filtre particulaire)

1. Utiliser un estimateur préliminaire $\widehat{\theta}^{(0)} = (\widehat{\theta}_1^{(0)'}, \dots, \widehat{\theta}_S^{(0)'})'$ du vecteur θ , où

$$\widehat{\theta}_s^{(0)} = \left(\widehat{\alpha}_s^{(0)}, \left(\text{vech} \left(\widehat{\Phi}_s^{(0)} \right) \right)', \left(\text{vech} \left(\widehat{\Sigma}_e^{(s,0)} \right) \right)', \left(\text{vech} \left(\widehat{\Sigma}_\eta^{(s,0)} \right) \right)' \right)', \quad s = \overline{1, S}.$$

2. Exécuter l'algorithme EM comme suit. A la $i^{\text{ème}}$ itération :

(a) Etape E :

i. Appliquer le filtre particulaire (Algorithme 3.2) pour obtenir $\left\{ f_t^{(j)}, j = \overline{1, M} \right\}$ afin d'approximer $p \left(X_t | Y_t, \theta^{(i-1)} \right)$.

ii. Appliquer le lissage particulaire (Algorithme 3.4) pour obtenir $\left\{ s_t^{(j)}, j = \overline{1, M} \right\}$ afin d'approximer $p \left(X_t | Y_n, \theta^{(i-1)} \right)$ et de calculer les quantités $\widehat{X}_t^{(n)}, \widehat{P}_t^{(n)}, \widehat{P}_{t,t-1}^{(n)}, \widehat{A}_t$ et \widehat{B}_t .

(b) Etape M : déterminer, pour tout $s = \overline{1, S}$, la valeur de $\widehat{\alpha}_s^{(i)}$, $\text{vech} \left(\widehat{\Phi}_s^{(i)} \right)$, $\text{vech} \left(\widehat{\Sigma}_e^{(s,i)} \right)$ et $\text{vech} \left(\widehat{\Sigma}_\eta^{(s,i)} \right)$, à partir de (3.7).

3.5 Application

Cette dernière section fournit une application empirique. Plus précisément, nous décrivons d'abord nos données et fournissons une analyse préliminaire afin de justifier l'utilisation de modèle périodique et multivarié. Nous appliquons donc l'analyse théorique développée dans ce chapitre pour fournir une modélisation multivariée du prix du pétrole et de trois séries temporelles de taux de change.

3.5.1 Données et analyse préliminaire

Nous modélisons quatre séries chronologiques : la série journalière des prix du pétrole Saharan Blend (SB)¹ et trois séries chronologiques quotidiennes de taux de change. Il s'agit de la Monnaie Européenne face au Dollar américain (EUR/USD)², la Monnaie Européenne face au Dinar Algérien (EUR/DZD) et le Dollar Américain face au Dinar Algérien (USD/DZD)³. Nous avons d'abord supprimé tous les jours où le marché SB était fermé. La période d'analyse s'étend du 3 janvier 2005 au 31 décembre 2015. L'échantillon final comprend 2836 observations pour chaque série.

Dans cette étude empirique, le modèle est structuré comme suit : nous collectons le prix du pétrole

¹Source de données : Ministère de l'Energie Algérien (Algérie).

²Source de données : EFX, Thomson Reuters.

³Source de données : Banque d'Algérie.

du Sahara Blend et les trois taux de change à l'instant t dans le vecteur

$$Y_t = \begin{pmatrix} Y_{t,1} \\ Y_{t,2} \\ Y_{t,3} \\ Y_{t,4} \end{pmatrix} = \begin{pmatrix} SB_t \\ EUR/USD_t \\ EUR/DZD_t \\ USD/DZD_t \end{pmatrix},$$

et leurs séries des rendements logarithmiques correspondantes dans le vecteur $Z_t = (Z_{t,1}, Z_{t,2}, Z_{t,3}, Z_{t,4})'$, où $Z_{t,i} = \log(Y_{t,i}) - \log(Y_{t-1,i})$, pour $i = 1, 2, 3, 4$.

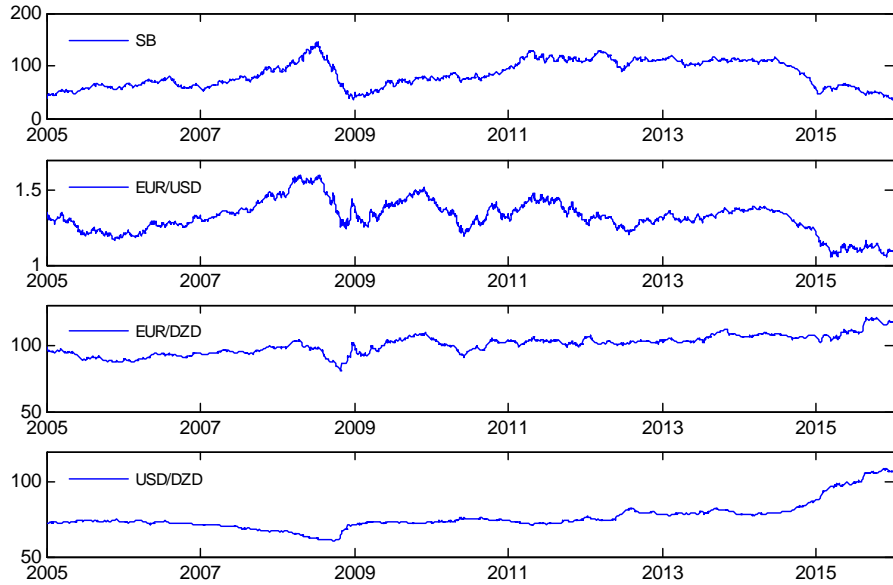


Figure 3.1 Représentation graphique des quatre séries brutes.

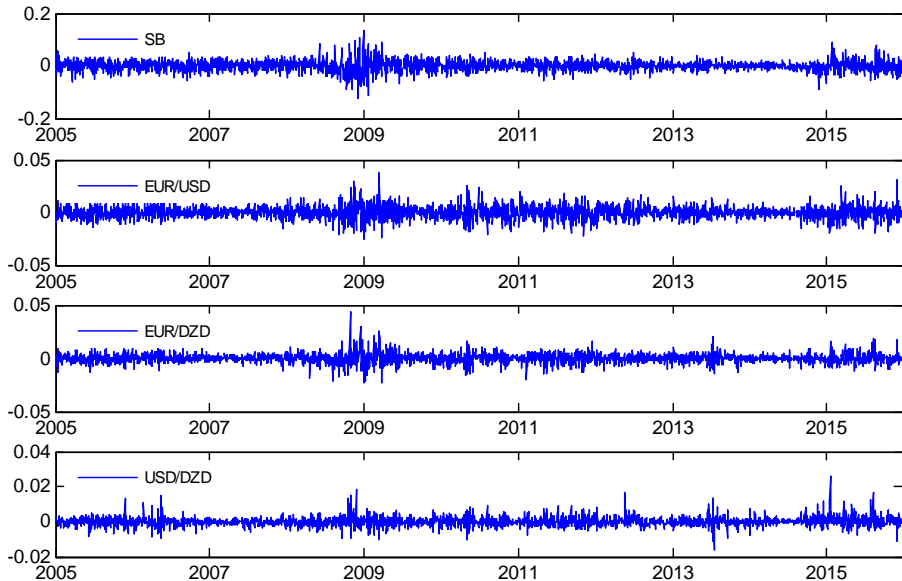


Figure 3.2 Représentation graphique des rendements logarithmiques.

La série brute $Y_{t,i}$ et la série des rendements $Z_{t,i}$ sont représentées respectivement dans les Figures 3.1 et 3.2. Il s'avère que toutes les séries brutes $Y_{t,i}$, $i = \overline{1,4}$, ne sont pas stationnaires, et en outre les séries des rendements semblent être stationnaires. Une propriété importante observée sur les quatre séries $Y_{t,i}$, $i = \overline{1,4}$, est le phénomène de regroupement des volatilités (volatility clustering). En effet, nous observons des sous-périodes de forte agitation suivies de sous-périodes beaucoup plus calmes (forte volatilité entre avril 2008 et juillet 2009 et après 2015, et faible volatilité entre décembre 2006 et mars 2008 et entre août 2012 et décembre 2014). Notons cependant l'extrême volatilité des différentes séries dans la période d'analyse, induite par la crise financière de l'automne 2008 et celle de 2015. De plus, durant toute la période de l'étude, la série des rendements SB est la plus volatile, suivie par les séries des rendements EUR/USD et EUR/DZD . La série des rendements USD/DZD est la plus stable.

En analysant quelques statistiques de base des quatre séries des rendements, présentées dans les Tableaux (3.1)-(3.4), il paraît évident que ces statistiques sont très similaires. Leurs écarts-types sont beaucoup plus importants que les moyennes en valeurs absolues, ce qui indique que les moyennes ne sont pas significativement différentes de zéro. De plus, leurs excès de Kurtosis sont significativement positifs ; par conséquent, ces séries ont des queues lourdes par rapport à la distribution normale. Le test de Jarque-Bera rejette la normalité des quatre distributions des rendements.

Tableau 3.1. Statistiques descriptives des rendements SB pour chaque jour de la semaine ainsi que toute la période d'étude

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	tous les jours
# of obs.	567	570	572	569	557	2835
Mean	-0.0014	0.0000	0.0005	0.0011	-0.0004	-2.29×10^{-5}
Median	0.0000	0.0000	0.0011	0.0020	-0.0003	9.32×10^{-5}
Maximum	0.1367	0.0633	0.0850	0.1067	0.0825	0.1367
Minimum	-0.0777	-0.1119	-0.0812	-0.0909	-0.1244	-0.1244
Std. Dev.	0.0212	0.0199	0.0182	0.0211	0.0193	0.0110
Skewness	0.7466	-0.4172	-0.0594	0.0078	-0.3224	0.0276
Kurtosis	8.5317	5.4482	5.0156	5.9985	7.9584	6.7408
Jarque-Bera	775.5827	158.8808	97.1668	213.1664	580.2407	1653.3850

Tableau 3.2. Statistiques descriptives des rendements EUR/USD pour chaque jour de la semaine ainsi que toute la période d'étude

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	tous les jours
# of obs.	567	570	572	569	557	2835
Mean	-0.0005	-0.0003	0.0000	0.0002	0.0001	-7.56×10^{-5}
Median	-0.0001	-0.0001	0.0000	0.0001	0.0002	0.0000
Maximum	0.0236	0.0264	0.0272	0.0381	0.0301	0.0381
Minimum	-0.0332	-0.0213	-0.0206	-0.0254	-0.0212	-0.0332
Std. Dev.	0.0072	0.0058	0.0063	0.0065	0.0067	0.0065
Skewness	-0.2244	0.2663	0.1105	0.2828	0.1631	0.0833
Kurtosis	4.3099	5.6897	4.5673	5.9422	4.5940	5.0206
Jarque-Bera	45.2972	178.5543	59.7092	212.8093	61.4376	485.5663

Tableau 3.3. Statistiques descriptives des rendements EUR/DZD pour chaque jour de la semaine ainsi que toute la période d'étude

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	tous les jours
# of obs.	567	570	572	569	557	2835
Mean	-0.0002	-0.0002	0.0003	0.0003	0.0001	6.32×10^{-5}
Median	-0.0001	0.0001	0.0002	0.0004	0.0000	0.0001
Maximum	0.0209	0.0170	0.0313	0.0497	0.0226	0.0497
Minimum	-0.0233	-0.0230	-0.0158	-0.0195	-0.0220	-0.0233
Std. Dev.	0.0049	0.0041	0.0044	0.0049	0.0048	0.0046
Skewness	-0.0819	-0.4649	0.8022	2.2173	0.1346	0.6007
Kurtosis	6.2042	6.6266	8.9065	25.1483	6.8762	11.6985
Jarque-Bera	243.1895	332.8976	892.8278	12096.3000	350.3907	9108.4040

Tableau 3.4. Statistiques descriptives des rendements USD/DZD pour chaque jour de la semaine ainsi que toute la période d'étude

	Lundi	Mardi	Mercredi	Jeudi	Vendredi	tous les jours
# of obs.	567	570	572	569	557	2835
Mean	0.0003	0.0000	0.0000	0.0001	0.0003	0.0001
Median	0.0001	0.0000	-0.0001	-0.0001	0.0001	-1.89×10^{-5}
Maximum	0.0282	0.0094	0.0138	0.0181	0.0335	0.0335
Minimum	-0.0092	-0.0113	-0.0070	-0.0100	-0.0132	-0.0132
Std. Dev.	0.0030	0.0023	0.0024	0.0028	0.0033	0.0028
Skewness	1.8854	0.0974	1.0514	1.0981	2.8623	1.8005
Kurtosis	17.9821	6.3905	8.5105	9.5237	27.1108	18.9315
Jarque-Bera	5638.8570	273.9248	829.1143	1123.3530	14252.2300	31513.4400

Le Tableau 3.5 présente la matrice de corrélation de l'échantillon pour toute la période d'étude. Les coefficients de corrélation varient de -0.5591 à 0.6720 . La corrélation négative la plus forte est pour la paire (EUR/USD , USD/DZD) tandis que la corrélation positive la plus forte est pour la paire (EUR/USD , EUR/DZD). Le SB est le moins corrélé avec tous les taux de change. Pour cela, nous suggérons d'analyser ces quatre séries en utilisant une modélisation multivariée qui peut inclure non seulement la structure de dépendance de chaque série de rendements, mais aussi l'interdépendance entre les différentes séries.

Table 3.5. Matrice de corrélation des rendements

Rendements	SB	EUR/USD	EUR/DZD	USD/DZD
SB	1	0.0491	0.0926	-0.1260
EUR/USD		1	0.6720	-0.5591
EUR/DZD			1	-0.4003
USD/DZD				1

Par ailleurs, les Tableaux 3.1-3.4 présentent quelques statistiques de base des quatre séries des rendements pour chaque jour de la semaine. Les résultats indiquent que les plus bas ($-0,0132$) (resp. -0.0233 , $-0,0332$, $-0,1244$) et les plus hauts (0.0335) (resp. 0.0497 , 0.0381 , 0.1367) rendements sont observés vendredi (vendredi et jeudi, lundi et jeudi, vendredi et lundi). Pour la période analysée, les statistiques descriptives par jour enregistrent le plus petit écart-type de $0,0023$ (resp. 0.0041 , 0.0058 ,

0.0182) le mardi (mardi, mardi, mercredi) pour la série des rendements USD/DZD (resp. EUR/DZD , EUR/USD , SB), tandis que l'écart-type le plus élevé de 0.0033 (resp. 0.0049, 0.0072, 0.0212) est enregistré le vendredi (resp. lundi, jeudi et lundi). La skewness pour tous les jours de la semaine et pour toutes les séries des rendements est positive, sauf les lundi et mardi pour EUR/DZD , le lundi pour EUR/USD et les mardi, mercredi et vendredi pour SB . De plus, l'une des caractéristiques qui ressort le plus clairement des Tableaux 3.1-3.4 est que la Kurtosis pour chaque jour de la semaine est supérieure à 3 d'une manière significative. Cela reflète le fait que les queues des distributions de toutes les séries des rendements journalières analysées sont plus épaisses que celles de la distribution normale (à décroissance plus lente que $\exp(-x^2/2)$).

Tableau 3.6. Autocorrélations périodique des séries des rendements

Rendements	Lundi	Mardi	Mercredi	Jeudi	Vendredi
SB	0.0565	-0.0383	0.0252	0.0596	0.0241
EUR/USD	0.0272	-0.1629	0.0182	-0.0006	-0.0339
EUR/DZD	0.0424	-0.0158	0.0231	0.2094	0.0877
USD/DZD	0.1617	-0.0617	-0.0685	0.1639	0.1728

Ces statistiques descriptives des quatre séries pour chaque jour de la semaine ne sont pas surprenantes car elles sont généralement observées dans les séries chronologiques financières où les corrélations entre les rendements quotidiens et la volatilité affichent des effets de jour de la semaine. Une telle propriété peut être utile car parfois on peut s'attendre à ce que les agents économiques aient un comportement différent à chaque jour de la semaine. En effet, d'après les autocorrélations quotidiennes de premier ordre pour nos quatre séries des rendements (voir Tableau 3.6), il est évident que les rendements du lundi sont positivement corrélés avec ceux des vendredis précédents, alors que les rendements du mardi sont corrélés négativement avec ceux du lundi. Pour cela, nous considérons une formulation périodique afin de prendre en charge cette caractéristique importante dans la modélisation et l'analyse de nos données.

3.5.2 Résultats empiriques

Nous proposons ci-après un modèle périodique qui permet de décrire *l'effet du jour de la semaine* dans les séries journalières de taux de change et du prix de pétrole. Une telle formulation 5-période semble bien adaptée pour expliquer la présence (ou non) de la saisonnalité. Plus précisément, nous considérons une formulation 5-périodique, où les paramètres varient selon le jour de la semaine ($s = 1, 2, 3, 4$ et 5)

$$s = \begin{cases} 1 & \text{si le jour correspondant est le lundi,} \\ 2 & \text{si le jour correspondant est le mardi,} \\ \vdots & \\ 5 & \text{si le jour correspondant est le vendredi.} \end{cases}$$

D'après le Tableau 3.6, il est clair que la série $Z_t, t = \overline{1, 2835}$, présente une périodicité dans la moyenne conditionnelle. Afin de capturer une telle périodicité, ce qui n'est pas possible par l'utilisation d'un modèle *PVAR-SV* pur, nous avons appliqué à Z_t un filtre autorégressif périodique multivarié (*PVAR*) et nous avons considéré la série des résidus comme la série sous-jacente de notre modélisation *PVAR-SV*. Le Tableau 3.7 présente les résultats de l'application de ce filtre.

Tableau 3.7. Le filtre *PVAR* appliqué sur $Z_t : Z_{s+5\tau} = \mu_s + \Psi_{s,1}Z_{s+5\tau-1} + \Psi_{s,2}Z_{s+5\tau-2} + \epsilon_{s+5\tau}$.

	μ_s	$\Psi_{s,1}$	$\Psi_{s,2}$
$s = 1$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0.0003 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.0504 & -0.3832 & 0.5118 & -0.5567 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1306 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.4004 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1020 \end{pmatrix}$
$s = 2$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.0495 & -0.6245 & 0.5911 & -0.5205 \\ 0 & 0 & 0 & 0 \\ -0.0162 & 0.0437 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -0.2530 & 0.2387 & -0.4112 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$s = 3$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.0495 & -0.5059 & 0.6890 & -0.3989 \\ 0.0217 & 0 & 0 & 0 \\ -0.0118 & 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.0309 & -0.2503 & 0.3707 & 0 \\ 0.0219 & 0 & 0.1018 & 0.3009 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$s = 4$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.0496 & -0.4237 & 0.5522 & -0.4355 \\ 0.0489 & 0 & 0.2837 & 0.4225 \\ -0.0155 & 0 & 0.1460 & 0.2786 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -0.1685 & 0.3661 & 0 \\ 0 & 0 & 0.1323 & 0.2303 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$s = 5$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5124 & -0.4541 & 0 \\ 0.0566 & -0.4402 & 0.3986 & -0.6652 \\ 0.0236 & 0.1061 & 0 & 0.2977 \\ 0 & 0 & 0.3098 & 0.4023 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -0.1508 & 0.3265 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -0.0876 & -0.1330 \end{pmatrix}$

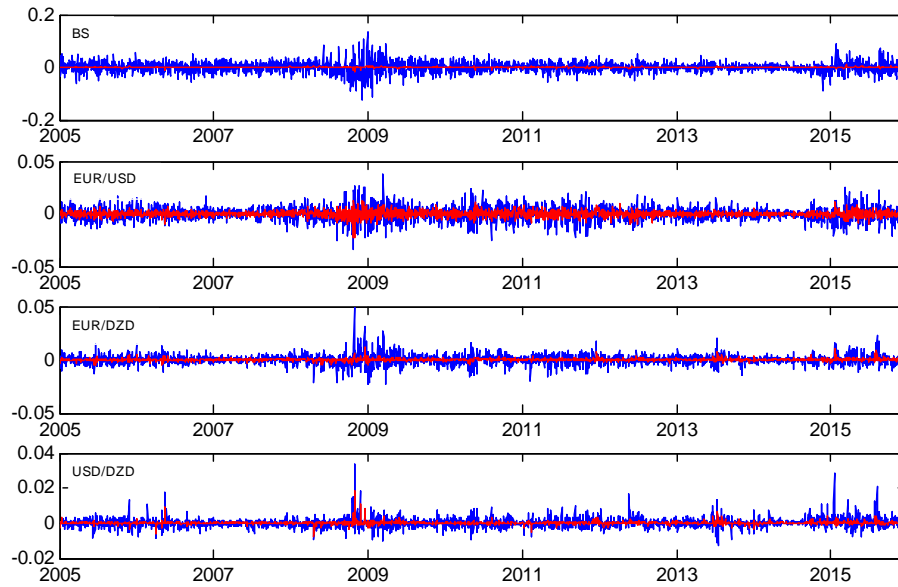


Figure 3.3 Les rendements logarithmiques et les séries ajustées à partir du modèle $PVAR_5(2)$ estimé.

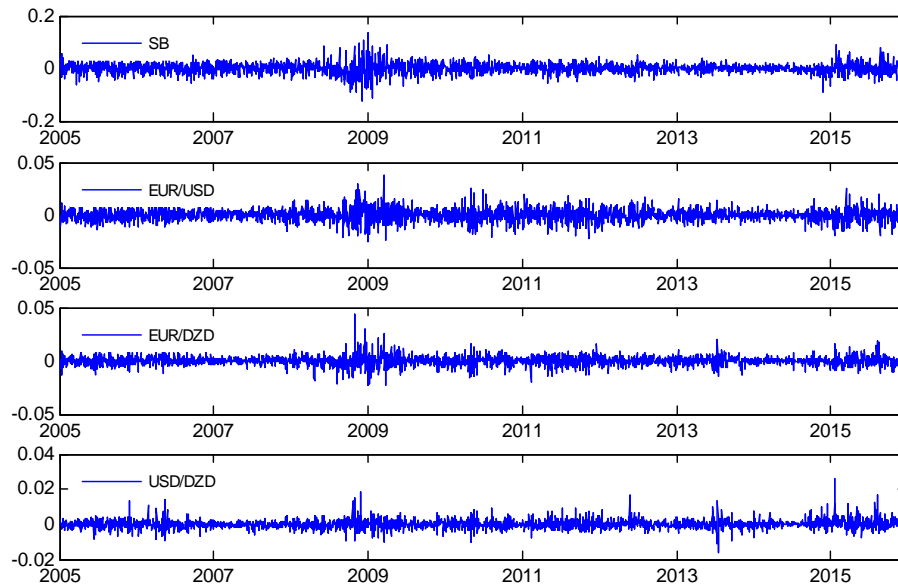


Figure 3.4 Les séries résiduelles calculées à partir du modèle $PVAR_5(2)$ estimé.

Les graphes des séries réelles $Z_{t,i}, i = \overline{1,4}$, ajustées et les valeurs résiduelles calculées à partir du modèle $PVAR_5(2)$ estimé sont données dans les Figures 3.3-3.4. Dans un second temps, nous utilisons ces résidus pour l'estimation de la volatilité.

Table 3.8. Estimation des paramètres du modèle *PVAR-SV* 5-periodique $\mathbf{e}_{s+5\tau} : \mathbf{e}_{s+5\tau} = V_{s+5\tau}^{1/2} \boldsymbol{\eta}_{s+5\tau}$, $X_{s+5\tau} = \boldsymbol{\alpha}_s + \Phi_s X_{s+5\tau-1} + \mathbf{e}_{s+5\tau}$, $V_t = \text{diag}(\exp X_t)$, $\boldsymbol{\eta}_{s+5\tau} \sim \mathcal{N}\left(0_{4 \times 1}, \Sigma^{(s)}\right)$ et $\mathbf{e}_{s+5\tau} \sim \mathcal{N}\left(0_{4 \times 1}, \Sigma^{(s)}\right)$

	$\boldsymbol{\alpha}_s$	Φ_s
$s = 1$	$\left\{ \begin{array}{l} -0.20262 \\ 0.01157 \\ 0.00686 \\ 0.00704 \end{array} \right\}$	$\left\{ \begin{array}{l} 0.98040 \\ 0.00043 \quad 0.99685 \\ 0.00024 \quad -0.00058 \quad 0.99705 \\ 0.00038 \quad 0.00067 \quad -0.00015 \quad 0.99605 \end{array} \right\}$
$s = 2$	$\left\{ \begin{array}{l} -1.45498 \\ -1.20156 \\ -0.86349 \\ -0.61738 \end{array} \right\}$	$\left\{ \begin{array}{l} 0.83309 \\ 0.00423 \quad 0.95710 \\ 0.03949 \quad 0.02311 \quad 0.94923 \\ 0.00272 \quad 0.05241 \quad 0.04000 \quad 0.94269 \end{array} \right\}$
$s = 3$	$\left\{ \begin{array}{l} 3.91287 \\ -0.00441 \\ -0.05308 \\ -0.01491 \end{array} \right\}$	$\left\{ \begin{array}{l} 1.51103 \\ 0.00338 \quad 0.91994 \\ 0.00181 \quad -0.00097 \quad 0.91795 \\ 0.00031 \quad 0.00238 \quad 0.00040 \quad 0.91900 \end{array} \right\}$
$s = 4$	$\left\{ \begin{array}{l} 0.69615 \\ -0.84999 \\ -0.55723 \\ -0.19088 \end{array} \right\}$	$\left\{ \begin{array}{l} 1.10199 \\ -0.01508 \quad 0.97555 \\ 0.03054 \quad 0.00381 \quad 0.96981 \\ 0.00130 \quad 0.03733 \quad 0.01311 \quad 0.98764 \end{array} \right\}$
$s = 5$	$\left\{ \begin{array}{l} -3.17158 \\ -0.50141 \\ -0.28571 \\ -0.07079 \end{array} \right\}$	$\left\{ \begin{array}{l} 0.55463 \\ -0.00491 \quad 0.91828 \\ 0.00438 \quad 0.01226 \quad 0.92180 \\ 0.00514 \quad 0.01030 \quad 0.01192 \quad 0.93396 \end{array} \right\}$

Table 3.8. (Suite.)

	$\Sigma_{\eta}^{(s)}$	$\Sigma_{\mathbf{e}}^{(s)} \times 10^{-3}$
$s = 1$	$\begin{pmatrix} 0.74094 \\ 0.01565 & 0.93385 \\ 0.08498 & 0.62127 & 0.90496 \\ -0.08795 & -0.63367 & -0.51178 & 0.91656 \\ 0.89185 \\ 0.05403 & 1.09090 \\ 0.09130 & 0.77576 & 1.29935 \\ -0.13203 & -0.69250 & -0.80344 & 1.30805 \\ 0.89839 \\ 0.07722 & 0.68868 \\ 0.10155 & 0.49622 & 0.69023 \\ -0.10294 & -0.39278 & -0.34332 & 0.66463 \\ 1.15961 \\ 0.02547 & 1.17425 \\ 0.03613 & 0.78995 & 1.10038 \\ -0.10227 & -0.82562 & -0.71740 & 1.26023 \\ 0.64827 \\ 0.06520 & 0.94032 \\ 0.07977 & 0.70570 & 0.97110 \\ -0.10650 & -0.64738 & -0.57424 & 1.06039 \end{pmatrix}$	$\begin{pmatrix} 0.00001 \\ 0.00001 & 0.09361 \\ 0.00001 & -0.00100 & 0.09663 \\ 0.00001 & -0.00029 & 0.00194 & 0.09346 \\ 0.00001 \\ 0.00229 & 131.77359 \\ -0.00209 & 12.19494 & 124.73988 \\ 0.00239 & 10.40001 & 10.84510 & 127.80300 \\ 0.54644 \\ 0.02714 & 2.54788 \\ 0.02448 & -0.06642 & 2.54067 \\ 0.02688 & -0.06561 & -0.02913 & 2.49501 \\ 224.15549 \\ 7.38600 & 73.34013 \\ -0.13731 & 7.59376 & 66.03193 \\ 6.12611 & 4.87846 & 2.96769 & 68.00796 \\ 57.78523 \\ 1.85298 & 38.40728 \\ 2.47936 & 2.24313 & 37.65754 \\ 0.28442 & -0.23460 & 1.51726 & 35.59162 \end{pmatrix}$
$s = 2$		
$s = 3$		
$s = 4$		
$s = 5$		

Les paramètres estimés du modèle $PVAR-SV$ 5-périodique sont présentés dans le Tableau 3.8. Nous remarquons que toutes les séries des rendements examinées présentent des volatilités très persistantes. Le coefficient de persistance varie entre $\prod_{s=1}^S \phi_{i,i}^{(s)} = 0.7960$ pour le USD/DZD à 0.7543 pour le SB . Cela, nous permet de conclure que le modèle estimé $PVAR-SV$ est périodiquement stationnaire puisque la quantité caractérisant les processus estimés, à savoir $\rho \left(\prod_{s=1}^S \Phi_s \right)$ définie dans (3.4), est plus petite que l'unité ($\rho \left(\prod_{s=1}^S \Phi_s \right) = 0.8009$). Par contre, lorsque nous analysons le modèle estimé période par période, nous remarquons que les modèles correspondants à mercredi et à jeudi ne sont pas stationnaires. En effet, nous avons $\rho(\Phi_s)$ égal, respectivement, à 0.9971 , 0.9571 , 1.5110 , 1.1020 et 0.9340 pour $s = 1, 2, 3, 4$ et 5 .

Ces résultats sont cohérents avec le fait que les volatilités des taux de change et des prix du pétrole sont souvent regroupés (plus $\prod_{s=1}^S \phi_{i,i}^{(s)}$ est grand plus l'effet sur la volatilité d'une grande valeur de $|\eta_{t,i}|$ disparaît lentement). Dans tous les cas, les coefficients de transmission ($\phi_{i,j}^{(s)}$) sont relativement faibles et il existe quand même une interaction de volatilité dans les données quotidiennes. Par exemple, à partir du modèle estimé, les informations sur le marché du pétrole sont rapidement transmises et incorporées dans les taux de change EUR/USD , EUR/DZD et USD/DZD . Cela pourrait suggérer que la volatilité du marché pétrolier a des impacts significatifs sur le marché des taux de change, et en particulier sur le Dinar Algérien.

Un autre phénomène intéressant apparaît dans notre analyse empirique, il s'agit du fait que la corrélation périodique des chocs $\frac{\Sigma_{\eta}^{(s)}(i,j)}{\sqrt{\Sigma_{\eta}^{(s)}(i,i)}\sqrt{\Sigma_{\eta}^{(s)}(j,j)}}$ ($s = \overline{1,5}$ et $i, j = \overline{1,4}$) est significativement négative entre USD/DZD et les autres séries. Cependant cette corrélation périodique est significativement positive dans tous les autres cas (entre EUR/USD et SB et entre EUR/DZD et SB et EUR/USD).

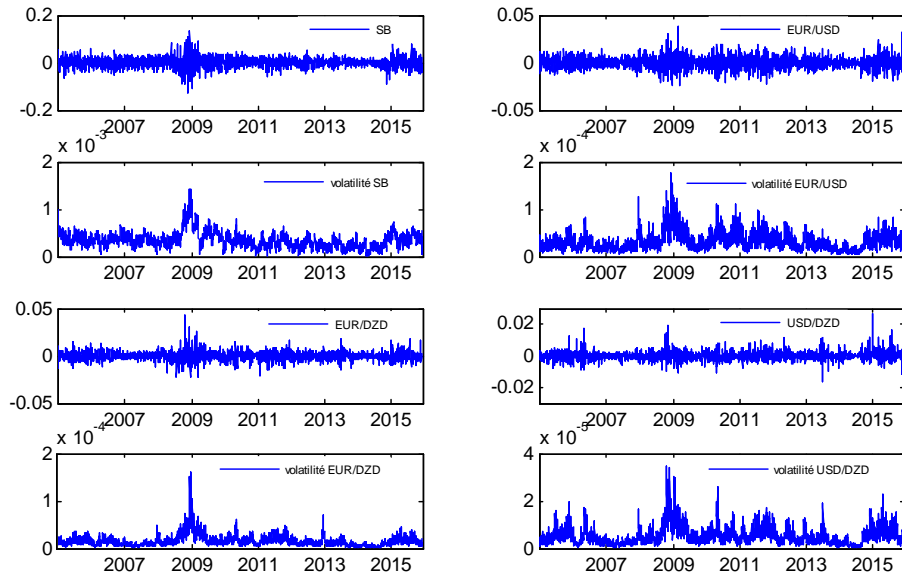


Figure 3.5 Les volatilités estimées à partir du modèle $PVAR-SV$ 5-périodique.

La Figure 3.5 présente les volatilités estimées à partir du modèle $PVAR-SV$ 5-périodique de SB , EUR/USD , EUR/DZD et USD/DZD . Nous pouvons constater que les volatilités estimées reflètent

bien la variation des résidus. De plus il s'avère, à la lecture du Tableau 3.9, que les intervalles de prédiction asymptotiques établis à partir de la modélisation *PVAR-SV*, sont proches des intervalles de prédiction théoriques. Ces résultats montrent que le modèle *PVAR-SV* ajusté aux résidus calculés à partir de la modélisation *PVAR* de nos séries journalières des rendements logarithmiques, semble être bien précis et présente une bonne performance de prévision de la volatilité.

Tableau 3.9. Intervalles de prédiction asymptotiques à l'horizon 1, à $(1 - \alpha) 100\%$.

	50%	60%	70%	80%	90%	95%	99%
Saharan Blend oil prices	53.81	62.50	71.61	79.80	87.96	92.69	96.96
Euro/Dollar Américain	53.25	62.96	72.60	81.04	89.65	93.33	97.78
Euro/Dinar Algérien	52.37	61.62	70.37	79.41	87.71	92.02	96.36
Dollar Américain/Dinar Algérien	53.50	62.61	71.68	79.31	87.25	91.56	95.97

Chapitre 4

Mélange de Modèles à Volatilité Stochastique

4.1 Introduction

Un modèle de mélange fini de lois de probabilité est une combinaison convexe de deux ou plusieurs fonctions de densité de probabilité. En combinant les propriétés individuelles des densités de probabilité, les modèles de mélange sont capables d'approximer toute distribution arbitraire (McLachlan and Peel, 2004). Par conséquent, les modèles de mélange fini sont un outil puissant et flexible et un moyen d'obtenir des distributions relativement complexes, tout en s'appuyant sur des lois très simples, comme les lois normales.

Les modèles de mélange fini ont été utilisés pendant plus de 100 ans. Il s'agit, en effet, d'une méthode extrêmement souple permettant de modéliser des phénomènes divers et variés. Ils ont, en pratique, reçu une attention croissante au fil des ans, tant d'un point de vue théorique que pratique. Ils ont même connu un véritable regain de popularité au cours des dernières décennies en raison de l'énorme augmentation de la puissance de calcul disponible. L'étendue et le potentiel des applications de modèles de mélanges finis, qui vont de la biologie et la médecine à la physique, à l'économie et la commercialisation, au marketing, et aux sciences de l'ingénieur et la reconnaissance d'images, ont été considérablement élargis. Des logiciels comme `mixmod1` (Biernacki et al., 2006) mettent à disposition ces méthodes sous forme de codes performants et portables et aident largement à leur diffusion.

Bien que le mélange de modèles statistiques soit apparu en 1894 avec les recherches de Pearson, il n'a été introduit qu'en 1978 dans l'analyse des séries chronologiques avec les modèles à seuil. Le mélange de modèles de séries chronologiques consiste généralement en un ensemble de modèles et un processus de transition latent.

Selon la structure de cette transition, nous pouvons classer les modèles à changement de régime en trois catégories : modèles à seuil (Tong, 1978), modèles à changement de régime Markovien (Hamilton, 1989) et modèles à changement de régimes *i.i.d.* (Wong et Li, 2000).

Des recherches récentes, utilisant des modèles de type-*ARCH*, ont été développées dans ce sens. En effet, Wong et Li (2001) ont proposé d'ajouter plus de flexibilité à la famille de mélange de modèles autorégressifs (*MAR*), en introduisant un mélange de modèles autorégressifs à erreurs *ARCH* (*MAR-ARCH*). Plusieurs extensions du mélange de modèles *MARCH* (Wong et Li, 2001) ont été proposées dans la littérature des séries chronologiques. Citons par exemple, le mélange de modèles *GARCH* (Haas et al., 2012 et Zhang et al., 2006), le mélange de modèles *GARCH* multivariés (Bauwens et al., 2006b), mélange de modèles *GARCH* à seuil (Giannikis et al., 2008), le mélange de modèles *ARCH* périodiques (Bentarzi et Hamdi, 2008a), le mélange de modèles *AR-ARCH* périodiques (Bentarzi et Hamdi, 2008b), le mélange de modèles *GARCH* périodique (Hamdi et Souam, 2013, 2017) et bien d'autres.

Cependant, en comparaison avec la littérature relative aux mélanges de modèles type-*ARCH*, la littérature sur les mélanges *SV* est trop maigre. Citons par exemple les travaux de Chen et al. (2008), Elliott et al. (2011) et Mao et al. (2017) sur les modèles *SV* à seuil. Nous citons également les travaux de So et al. (1998) et Carvalho (2007) sur les modèles *SV* à changement de régimes Markovien. D'autres travaux ont été consacrés au mélange *i.i.d.* de modèles *SV*. En ce qui concerne cette dernière catégorie de modèles, Kim et Stoffer (2008) ont proposé d'introduire un mélange de deux lois normales pour approximer la distribution de l'erreur de l'équation d'observation linéarisée. Aucune hypothèse préalable n'est exigée pour la distribution de l'erreur d'observation. Xu et Knight (2013) ont proposé un modèle plus général. Ils ont présenté un mélange de plusieurs lois normales bi-variées pour le modèle *SV* linéarisé dans le but de capturer l'effet de levier. Pour notre part, nous proposons une autre généralisation qui consiste en un mélange de K composantes *SV*, et qui peut être définie à travers la distribution conditionnelle du processus sous-jacent, compte tenu de K volatilités parallèles.

Ce chapitre est organisé comme suit. Nous présentons dans la Section 4.2, la définition générale du mélange de modèles pour nous restreindre rapidement au cas gaussien et aux mélanges *i.i.d.* Nous donnons comme exemples le modèle *MAR-ARCH* de Wong et Li (2001) et quelques mélanges de modèle *SV* existants dans la littérature. La Section 4.3 est consacrée à la présentation d'une nouvelle classe de Mélange de modèles à Volatilité Stochastique (*M-SV*). Dans la Section 4.4, nous étudions les conditions de stationnarité et d'ergodicité. Le problème de l'estimation des paramètres par la méthode *ML* basée sur le filtrage et le lissage particulaire est traité dans la Section 4.5. Enfin, la bonne performance de la méthode *ML* est montrée par une étude de simulation.

4.2 Mélange de modèles gaussiens

4.2.1 Mélange de lois de probabilité

Définition 4.1 Soient Y et Z deux variables aléatoires et soient $\phi_{Y|Z=z}$ la densité de la distribution de Y conditionnellement à $Z = z$ et H est la fonction de répartition de Z . La forme générale de la densité de mélange s'écrit comme suit

$$f(y) = \int \phi_{Y|Z=z}(y) dH(z).$$

Quand Z ne prend qu'un ensemble fini de valeurs $\{1, \dots, K\}$, alors la distribution de Z est discrète et affecte une probabilité positive à chacune de ces valeurs. L'intégrale est alors remplacée par une somme pour donner une densité de mélange fini. Cette densité est définie par

$$f(y) = \sum_{k=1}^K \lambda_k \phi_k(y),$$

où $\phi_k(y) = \phi_{Y|Z=k}(y)$ et $\lambda_k = \mathbb{P}(Z = k)$. Ces probabilités λ_k sont appelées proportions du mélange et sont assujetties aux deux contraintes suivantes

$$\lambda_k > 0, \text{ pour tout } k \in \{1, \dots, K\} \text{ et } \sum_{k=1}^K \lambda_k = 1.$$

Par la propriété de convexité, étant donné que chaque ϕ_k définit une fonction de densité de probabilité, f sera également une fonction de densité de probabilité.

Comme cas particulier, nous trouvons le mélange gaussien qui est très classique. Dans la situation multivariée, la variable aléatoire Y est dans \mathbb{R}^n , la densité de la $k^{\text{ème}}$ composante du mélange ($k = \overline{1, K}$) s'écrit comme suit

$$\phi_k(y|\theta_k) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{|\Sigma_k|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (y - \mu_k)' \Sigma_k^{-1} (y - \mu_k) \right\},$$

avec $\theta_k = (\mu_k', \text{vec}'(\Sigma_k))'$, μ_k est la moyenne de la $k^{\text{ème}}$ composante du mélange et Σ_k sa matrice de variance-covariance.

Finalement, la densité de mélange gaussien s'exprime alors par

$$f(y|\underline{\theta}) = \sum_{k=1}^K \lambda_k \phi_k(y|\theta_k).$$

Cette dernière densité est paramétrée par le vecteur $\underline{\theta} = (\lambda', \theta_1', \dots, \theta_K')$, où λ est le vecteur des $K - 1$ premières proportions du mélange $(\lambda_1, \dots, \lambda_{K-1})'$.

L'importance du mélange de densités normales vient du fait que n'importe quelle densité peut être approchée par un mélange de lois normales (voir Marron et Wand, 1992). En effet, il est bien connu que les paramètres d'un mélange de deux lois gaussiennes univariées peuvent être choisis de telle sorte que la densité soit proche de celle d'une distribution log-normale. Pour illustrer ce point, Titterington et al. (1985) ont tracé la densité de mélange de deux lois gaussiennes univariées.

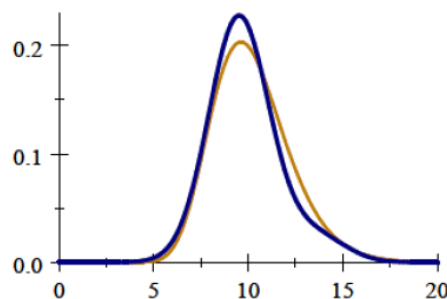


Figure 4.1. densité log-normale et mélange de deux densités normales
— $\log\mathcal{N}(\log(10), 0.04)$ — $0.9 \times \mathcal{N}(9.5, 2.5) + 0.1 \times \mathcal{N}(13.5, 2.5)$

Cette proximité entre la log-normale et la distribution du mélange de deux lois gaussiennes signifie qu'il est très difficile de faire la distinction entre elles dans la pratique. Cela a peu de conséquences, au moins dans un sens interprétatif, lorsque l'analyse se limite à l'obtention d'un modèle satisfaisant pour la distribution sous-jacente. Dans un sens calculatoire, nous avons besoin de certains efforts supplémentaires pour ajuster un modèle de mélange normal à cinq paramètres par rapport à un modèle log-normal à deux paramètres. Cependant, dans certaines situations, le choix entre les modèles log-normal et normal a beaucoup d'intérêt.

Marron et Wand (1992) exposent 15 exemples de mélanges de densités de lois normales choisies soigneusement pour qu'elles représentent de nombreux types de densités. Les cinq premières représentent de nombreux types différents de problèmes qui peuvent se présenter dans le cas des densités unimodales. Le reste des densités est multimodale (voir Figure 4.2).

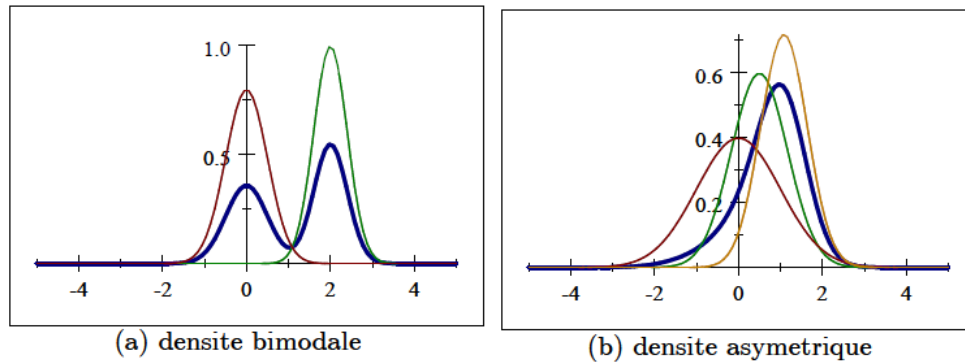


Figure 4.2. Exemples de mélanges de deux lois normales

(a) $\color{red}{\bullet} \mathcal{N}\left(0, \left(\frac{1}{2}\right)^2\right)$, $\color{green}{\bullet} \mathcal{N}\left(2, \left(\frac{2}{5}\right)^2\right)$ $\color{blue}{\bullet} 0.45 \times \mathcal{N}\left(0, \left(\frac{1}{2}\right)^2\right) + 0.55 \times \mathcal{N}\left(2, \left(\frac{2}{5}\right)^2\right)$

(b) $\color{red}{\bullet} \mathcal{N}(0, 1)$, $\color{green}{\bullet} \mathcal{N}\left(\frac{1}{2}, \left(\frac{2}{3}\right)^2\right)$ $\color{orange}{\bullet} \mathcal{N}\left(\frac{5}{9}, \left(\frac{13}{12}\right)^2\right)$ $\color{blue}{\bullet} 0.2 \times \mathcal{N}(0, 1) + 0.2 \times \mathcal{N}\left(\frac{1}{2}, \left(\frac{2}{3}\right)^2\right) + 0.6 \times \mathcal{N}\left(\frac{5}{9}, \left(\frac{13}{12}\right)^2\right)$

Remarque 4.1

- A partir des mélanges exposés par Marron et Wand (1992), la K -modalité n'implique pas qu'on a un mélange de K lois.
- L'unimodalité dans les histogrammes des données n'implique pas toujours que les données ont été échantillonnées à partir d'une seule distribution.

D'une façon générale, l'interprétation la plus directe des modèles de mélange de lois de probabilité est que la variable aléatoire X est générée à partir de K variables aléatoires distinctes X_k , $k = \overline{1, K}$. Chacune de ces variables X_k est représentée par sa densité ϕ_k , et λ_k représente la proportion d'observation de cette variable particulière. D'une manière plus simple, le modèle de mélange fini consiste à supposer que les données proviennent d'une source contenant plusieurs sous-populations. Chaque sous-population est modélisée de manière séparée. La population totale est un mélange de ces sous-populations. Le modèle résultant est un modèle de mélange fini.

Dans un modèle de mélange, plusieurs questions peuvent être posées concernant sa structure :

- Choix de la famille des densités : il existe plusieurs choix paramétriques pour la famille ϕ_k et même semi-paramétriques et non paramétriques (voir Bordes et al., 2006). On s'intéresse ici aux modèles paramétriques et particulièrement aux mélanges de modèles gaussiens.
- Détermination du nombre de composantes du mélange.
- Estimation des paramètres du mélange.

4.2.2 Quelques mélanges de modèles gaussiens

Mélange de modèles autorégressifs à erreurs ARCH

La classe des modèles de mélange autorégressifs à erreurs ARCH introduite et étudiée par Wong et Li (2001) est définie comme suit :

$$\left\{ \begin{array}{l} f(y_t | \mathcal{F}_{t-1}) = \sum_{l=1}^K \frac{\lambda_l}{\sqrt{h_t^{(l)}}} f^{(l)}\left(\frac{\varepsilon_{t,l}}{\sqrt{h_t^{(l)}}}\right) \\ \varepsilon_{t,k} = y_t - \phi_{0,k} - \sum_{i=1}^{p_k} \phi_{i,k} y_{t-i}, \\ h_t^{(k)} = \beta_{0,k} + \sum_{j=1}^{q_k} \beta_{j,k} \varepsilon_{t-j,k}^2, \end{array} \right. , t \in \mathbb{Z}, k = \overline{1, K},$$

où \mathcal{F}_{t-1} est le passé du processus jusqu'à l'instant $t-1$ et les paramètres $\beta_{j,t}$ sont tels que $\beta_{0,k} > 0$ et $\beta_{j,k} \geq 0$, pour tous $j = \overline{1, q}$, et $k = \overline{1, K}$. Si pour tous $i = \overline{1, p_k}$ et $k = \overline{1, K}$, $\phi_{i,k} = 0$, nous avons le modèle de mélange ARCH, noté MARCH ($K; q_1, q_2, \dots, q_K$).

Mélange de modèles à volatilité stochastique

Dans cette formulation, l'erreur de l'équation d'observation de la version linéarisée du modèle SV standard est remplacée par un mélange de lois normales. L'idée d'utiliser un mélange gaussien a été introduite par Kim et al. (1998). Ils l'utilisent pour des raisons de commodité dans la procédure d'échantillonnage de la méthode d'estimation MCMC. Ils ont proposé d'utiliser un mélange de sept lois normales (avec des paramètres particuliers) pour approcher la loi $\log \chi_1^2$. Omori et al. (2007) ont proposé d'étendre le mélange SV gaussien proposé par Kim et al. (1998) sans effet de levier à un mélange SV gaussien avec effet de levier. L'approche utilisée repose sur l'idée d'approximer la distribution conjointe du bruit de l'équation d'état et celui de l'équation d'observation par un mélange de distributions normales bivariées à dix composantes. Kim et Stoffer (2008) ont proposé un mélange de deux lois normales afin d'approximer la loi de l'erreur d'observation. Ils ont introduit le modèle suivant

$$\left\{ \begin{array}{l} h_t = \phi h_{t-1} + w_t, \\ y_t = h_{t-1} + v_t, \\ v_t = I_t z_{t,1} + (1 - I_t) z_{t,0} \end{array} \right. \quad t \in \mathbb{Z},$$

où $(z_{t,0})$ et $(z_{t,1})$ sont supposés *i.i.d.* de lois gaussiennes $z_{t,0} \sim \mathcal{N}(m_0, R_0)$, $z_{t,1} \sim \mathcal{N}(m_1, R_1)$, tels que $m_0 = \alpha - \pi\mu$, $m_1 = \alpha - (1 - \pi)\mu$. Ici les deux paramètres α et μ sont inconnus et la variable aléatoire I_t suit une loi de Bernoulli de paramètre π . Cependant, leur idée est différente de celle de Kim et al. (1998). Cette formulation permet d'estimer les paramètres des deux lois normales, en exploitant les

observations, même dans le cas où l'erreur d'observation¹ ne suit pas une loi $\log \chi_1^2$. Ils ont utilisé la méthode *ML* via l'algorithme *EM* combiné avec les filtres particulières.

Xu et Knight (2013) ont proposé une formulation plus générale contenant comme cas particuliers les deux modèles de Kim et Stoffer (2008) et de Omori et al. (2007). Ils ont proposé un mélange de K lois normales bivariées² pour le modèle *SV* linéarisé (*LSV-MN*). Ils ont proposé la formulation suivante

$$\begin{cases} x_t = \exp\left(\frac{h_t}{2}\right) e_t, \\ y_t = \log(x_t^2) = h_t + \varepsilon_t, \quad t \in \mathbb{Z}. \\ h_{t+1} = \lambda + \alpha h_t + \eta_t, \end{cases}$$

où

$$\begin{pmatrix} \varepsilon_t \\ \eta_t \end{pmatrix} \sim p_k \mathcal{N} \left(\begin{pmatrix} \mu_k \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_k^2 & \rho_k \sigma_k \sigma_\eta \\ \rho_k \sigma_k \sigma_\eta & \sigma_\eta^2 \end{pmatrix} \right), \quad k = \overline{1, K}.$$

Xu et Knight (2013) ont dérivé des expressions explicites des moments des puissances entières de $|x_t|$ et les moments du produit $|x_t|^m |x_t|^n$, $n, m \in \mathbb{N}$, ainsi que la fonction d'autocorrélation de $\{y_t; t \in \mathbb{Z}\}$. Une procédure d'estimation basée sur la fonction caractéristique empirique a été proposée pour l'estimation des paramètres.

4.3 Mélange de Modèles à Volatilité Stochastique

Considérons le processus $\{\varepsilon_t; t \in \mathbb{Z}\}$ satisfaisant (2.1). Il est important de rappeler ici que contrairement au cas de modèles *ARCH/GARCH*, dans l'approche *SV*, la variable h_t ne s'interprète plus comme la variance conditionnelle de ε_t à son passé, celle-ci n'a pas une forme explicite. En effet, dans le cas *SV*, nous avons

$$\text{var}(\varepsilon_t | \mathcal{F}_{t-1}) = \mathbb{E}(h_t | \mathcal{F}_{t-1}) \neq h_t,$$

où \mathcal{F}_t désigne la σ -algèbre engendrée par le passé du processus jusqu'à l'instant t .

Dans cette section, nous introduisons une nouvelle classe de modèles *SV* appelée Mélange de modèles à Volatilité Stochastique (*M-SV*). Ce modèle est défini à travers la distribution conditionnelle du processus sous-jacent, étant donné K volatilités parallèles, comme un mélange de K distributions paramétriques

$$\begin{cases} f(\varepsilon_t | Y_t) = \sum_{l=1}^K \frac{\lambda_l}{\sqrt{h_t^{(l)}}} f^{(l)}\left(\frac{\eta_t}{\sqrt{h_t^{(l)}}}\right), \\ \ln h_t^{(k)} = \alpha_k + \beta_k \ln h_{t-1} + e_{t,k}, \\ h_t := \text{Var}(\varepsilon_t | \mathcal{F}_{t-1}, Y_t) = \sum_{k=1}^K \lambda_k h_t^{(k)}, \end{cases}, \quad t \in \mathbb{Z}, \quad k = \overline{1, K}, \quad (4.1)$$

où $f^{(k)}$ et $f(\cdot | Y_t)$ sont, respectivement, une densité de probabilité centrée réduite et la fonction de densité de probabilité conditionnelle de ε_t , sachant le vecteur $Y_t = (Y_{t,1}, \dots, Y_{t,K})'$ des K log-volatilités parallèles à l'instant t . Notons que la $k^{\text{ème}}$ composante du vecteur Y_t est définie par $Y_{t,k} = \ln h_t^{(k)}$, pour tout $k = \overline{1, K}$, où $h_t^{(k)}$ est la volatilité du $k^{\text{ème}}$ régime. Le processus $\{e_t = (e_{t,1}, \dots, e_{t,K})', t \in \mathbb{Z}\}$ est un

¹L'approximation de densité dans Kim et al. (1998) et Omori et al. (2007) est basée sur l'hypothèse de normalité du terme d'innovation dans la formulation non-linéaire.

²Comme cela a été fait dans Omori et al. (2007).

bruit blanc de moyenne nulle et d'une structure de covariance définie par $\mathbb{E}(e_t e_{t+h}') = \delta_{h,0} \text{diag}(Q)$, où $Q = (Q_1, \dots, Q_K)' \in \mathbb{R}_+^{*K}$. Finalement, les constantes λ_k , $k = \overline{1, K}$, sont des réelles strictement positives tel que $\sum_{k=1}^K \lambda_k = 1$.

Notons ici que différemment aux modèles de mélange *ARCH/GARCH* et leurs extensions, on n'exprime pas le mélange *SV* à travers la fonction de répartition de ε_t sachant l'information disponible jusqu'à l'instant $t - 1$ mais sachant les K log-volatilités parallèles à l'instant t . Notons également que la volatilité du processus $\{\varepsilon_t; t \in \mathbb{Z}\}$ est donnée par

$$h_t = \text{Var}(\varepsilon_t | \mathcal{F}_{t-1}, Y_t) = \mathbb{E}(\varepsilon_t^2 | \mathcal{F}_{t-1}, Y_t) = \sum_{k=1}^K \lambda_k h_t^{(k)}, \quad t \in \mathbb{Z},$$

alors que la variance conditionnelle de ε_t , sachant l'information passée, est donnée par

$$\text{Var}(\varepsilon_t | \mathcal{F}_{t-1}) = \mathbb{E}(\varepsilon_t^2 | \mathcal{F}_{t-1}) = \sum_{k=1}^K \lambda_k \mathbb{E}(h_t^{(k)} | \mathcal{F}_{t-1}) = \mathbb{E}(h_t | \mathcal{F}_{t-1}), \quad t \in \mathbb{Z}.$$

Pour analyser la structure probabiliste du modèle *M-SV*, nous allons utiliser les hypothèses suivantes :

- les densités $f^{(k)}$, $k = \overline{1, K}$, sont strictement positives sur \mathbb{R} .
- introduisons le vecteur bivalent $Z_t = (Z_{t,1}, \dots, Z_{t,K})'$, dont chaque composante $Z_{t,k}$ est égale à 1 si l'observation ε_t est générée par la $k^{\text{ème}}$ composante de la distribution conditionnelle et 0 ailleurs. Ainsi, Z_t est distribuée selon une distribution multinomiale de paramètres $(1; \lambda_1, \lambda_2, \dots, \lambda_K)$, où λ_k est la probabilité que l'observation ε_t soit générée par la $k^{\text{ème}}$ composante de la distribution conditionnelle $f(\varepsilon_t | Y_t)$. De ce fait, nous avons

$$\mathbb{P}(Z_t = z_t) = \prod_{k=1}^K \lambda_k^{z_{t,k}},$$

et nous écrivons $Z_t \sim \text{Mult}_K(1, \lambda)$.

Le modèle (4.1) a été défini par sa distribution conditionnelle. Il est possible de le représenter par l'équation aux différences stochastique suivante

$$\begin{cases} \varepsilon_t = \left(\sum_{l=1}^K h_t^{(l)} \mathbf{1}_{(Z_{t,l}=1)} \right)^{\frac{1}{2}} \eta_t, \\ \ln h_t^{(k)} = \alpha_k + \beta_k \ln h_{t-1} + e_{t,k}, \\ h_t = \sum_{l=1}^K \lambda_l h_t^{(l)}, \end{cases}, \quad t \in \mathbb{Z}, \quad k = \overline{1, K}, \quad (4.2)$$

où $\{\eta_t; t \in \mathbb{Z}\}$ est une séquence de variables aléatoires *i.i.d.* centrées réduites. On peut exprimer $h_t^{(k)}$ d'une manière plus précise en utilisant la variable Z_t i.e. $h_t^{(k)} = \text{Var}(\varepsilon_t | \mathcal{F}_{t-1}, Y_t, Z_{t,k=1})$.

Notre idée est de construire un modèle avec K processus de volatilités parallèles, chacun d'eux a une dynamique *SV*. Mais dans cette formulation $\ln h_t^{(k)}$ ne suit pas un *AR(1)* et $\ln h_t$ non plus, par contre en regroupant les log-volatilités, $\ln h_t^{(k)}$, $k = 1, K$, dans le vecteur Y_t , on obtient une structure espace d'états simple et aisée d'estimation, où Y_t suit un modèle AutoRégressif Non-linéaire d'ordre un (*NAR(1)*). En effet, nous avons

$$Y_t = g(Y_{t-1}) + e_t, \quad (4.3)$$

où $g(Y_{t-1}) = \alpha + \beta \ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l})$, $\alpha = (\alpha^{(1)}, \dots, \alpha^{(K)})'$ et $\beta = (\beta_1, \dots, \beta_k)'$. La volatilité conditionnelle au $k^{\text{ème}}$ régime à l'instant t , $h_t^{(k)}$, dépend de toutes les volatilités conditionnelles aux autres régimes à l'instant $t - 1$

$$\ln h_t^{(k)} = \alpha_k + \beta_k \ln \left(\sum_{k=1}^K \lambda_k h_{t-1}^{(k)} \right) + e_{t,k},$$

et la volatilité du processus ε_t est une combinaison linéaire des volatilités conditionnelles aux régimes.

Il est à noter que cette formulation M - SV est différente des formulations de mélanges SV existants (Kim et al., 1998 ; Omori et al., 2007 ; et Kim et Stoffer, 2008 ; Xu et Knight, 2013).

4.4 Stationnarité et existence des moments

Rappelons qu'une fonction continue g est dite sous-linéaire s'il existe deux constantes positives a et b , et une norme $\|\cdot\|$ dans \mathbb{R}^K , telles que pour tout $x \in \mathbb{R}^K$,

$$\|g(x)\| \leq a \|x\| + b.$$

Dans notre cas la fonction $x \mapsto g(x) = \alpha + \beta \ln \sum_{l=1}^K \lambda_l \exp(x_l)$ est sous-linéaire. En effet, en utilisant la norme infinie

$$\begin{aligned} \|g(x)\|_\infty &= \left\| \alpha + \beta \ln \sum_{l=1}^K \lambda_l \exp(x_l) \right\|_\infty \leq \|\alpha\|_\infty + \|\beta\|_\infty \left\| \ln \sum_{l=1}^K \lambda_l \exp(x_l) \right\|_\infty \\ &= \max_{l=1,K} |\alpha^{(l)}| + \max_{l=1,K} |\beta^{(l)}| \left\| \ln \sum_{l=1}^K \lambda_l \exp(x_l) \right\|_\infty, \end{aligned}$$

et choisissons m de sorte que

$$|\lambda_m \exp(x_m)| = \max_{l=1,K} |\lambda_l \exp(x_l)|,$$

alors

$$\begin{aligned} \max_{l=1,K} |\alpha^{(l)}| + \max_{l=1,K} |\beta^{(l)}| \left\| \ln \sum_{l=1}^K \lambda_l \exp(x_l) \right\|_\infty &\leq \max_{l=1,K} |\alpha^{(l)}| + \max_{l=1,K} |\beta^{(l)}| \times \|\ln(K \lambda_m \exp(x_m))\|_\infty \\ &= \max_{l=1,K} |\alpha^{(l)}| + \max_{l=1,K} |\beta^{(l)}| \times \ln(K \lambda_m) \\ &\quad + \max_{l=1,K} |\beta^{(l)}| \times |x_m| \\ &\leq \max_{l=1,K} |\alpha^{(l)}| + \max_{l=1,K} |\beta^{(l)}| \times \ln \left(K \max_{l=1,K} \lambda_l \right) \\ &\quad + \max_{l=1,K} |\beta^{(l)}| \times \|x\|_\infty. \end{aligned}$$

Par suit, il est suffisant de choisir

$$a = \max_{l=1}^K |\beta^{(l)}| \quad \text{and} \quad b = \max_{l=1}^K |\alpha^{(l)}| + \max_{l=1}^K |\beta^{(l)}| \times \ln \left(K \max_{l=1}^K \lambda_l \right).$$

Proposition 4.1 (Boussaha et Hamdi, 2018) *Considérons le modèle sous-linéaire défini dans (4.2). Supposons que les conditions (1) et (2) ci-dessous sont vérifiées*

1. *il existe $\delta > 0$, tel que $\mathbb{E} \|e_1\|^\delta < \infty$.*
2. *la variable e_1 possède une densité par rapport à la mesure de Lebesgue et cette densité est strictement positive sur \mathbb{R}^K .*

Si de plus $\max_{l=\overline{1, K}} |\beta^{(l)}| < 1$, alors la chaîne de Markov (Y_t) est géométriquement ergodique.

Preuve. En utilisant des arguments similaires à ceux utilisés par Yao et Attali (2000) pour le cas des modèles *NAR* sous-linéaires, nous pouvons montrer l'ergodicité géométrique de (Y_t) . ■

La proposition suivante donne des conditions d'existence d'une solution stable d'ordre m au sens de Yao (Définition de la stabilité d'ordre S , 2000) de (4.2).

Proposition 4.2 (Boussaha et Hamdi, 2018) *Si les conditions (1), (2) de la Proposition 4.1 et $\mathbb{E} \|e_1\|^m < \infty$ sont vérifiées avec $m \geq 1$, alors la solution stationnaire possède un moment d'ordre m .*

4.5 Estimation des paramètres d'un modèle *M-SV*

4.5.1 Estimation des paramètres par la méthode *ML*

Notons par $X = (Y'_0, Y'_1, \dots, Y'_n, Z'_1, Z'_2, \dots, Z'_n, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$ et $(Y'_1, Y'_2, \dots, Y'_n, Z'_1, Z'_2, \dots, Z'_n)'$ le vecteur contenant respectivement les données complètes et les données inobservées. Soit $\underline{\theta} = (\theta', \lambda')'$, où $\theta = (\alpha', \beta', Q)'$ et $\lambda = (\lambda_1, \dots, \lambda_{K-1})'$. Pour une réalisation donnée $\underline{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ du modèle (4.1) et sous l'hypothèse supplémentaire que les $f^{(k)}$, $k = \overline{1, K}$, sont des distributions gaussiennes, la fonction de vraisemblance complète du vecteur des paramètres $\underline{\theta}$ peut être exprimée comme suit

$$\begin{aligned}
 L(\underline{\theta}; X) &= f(Y_0) \prod_{t=1}^n f(Y_t | \underline{Y}_{t-1}) \prod_{t=1}^n f(Z_t | \underline{Z}_{t-1}, \underline{Y}) f(\varepsilon_t | \underline{\varepsilon}_{t-1}, \underline{Y}, \underline{Z}) \\
 &= f(Y_0) \prod_{t=1}^n f(Y_t | Y_{t-1}) f(Z_t) f(\varepsilon_t | Y_t, Z_t) \\
 &= C \times f(Y_0) \prod_{t=1}^n \frac{1}{(2\pi) \left(\prod_{k=1}^K Q_k\right)^{1/2}} \exp\left(-\frac{(Y_t - g(Y_{t-1}))' (\text{diag}(Q))^{-1} (Y_t - g(Y_{t-1}))}{2}\right) \\
 &\quad \prod_{t=1}^n \prod_{k=1}^K \lambda_k^{Z_{t,k}} \left[\frac{1}{\sqrt{\exp Y_{t,k}}} \exp\left(-\frac{\varepsilon_t^2}{2 \exp Y_{t,k}}\right) \right]^{Z_{t,k}},
 \end{aligned}$$

où C est une constante indépendante du $\underline{\theta}$, et $\underline{Y}_t = (Y_0', Y_1', \dots, Y_t')$, $\underline{Z}_t = (Z_0', Z_1', \dots, Z_t')$ et $\underline{\varepsilon}_t = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_t)$. Donc,

$$\begin{aligned} -2\mathbf{L}(\underline{\theta}; X) &= -2 \log L(\underline{\theta}; X) \\ &= -2 \log f(Y_0) + \sum_{t=1}^n \sum_{k=1}^K \log Q_k + \sum_{t=1}^n (Y_t - g(Y_{t-1}))' \text{diag}(R) (Y_t - g(Y_{t-1})) \\ &\quad + \sum_{t=1}^n \sum_{k=1}^K Z_{t,k} \left(-2 \log \lambda_k + Y_{t,k} + \frac{\varepsilon_t^2}{\exp Y_{t,k}} \right) - 2 \log C. \end{aligned}$$

où $R = (\frac{1}{Q_1}, \dots, \frac{1}{Q_K})'$. La log-vraisemblance ne peut pas être maximisée directement, du fait que nous avons deux types de variables inobservables, la volatilité et la variable indicatrice. Dans cette situation, nous pouvons aussi utiliser l'algorithme *EM* comme dans les deux chapitres précédents.

L'espérance conditionnelle de la fonction log-vraisemblance est exprimée comme suit

$$\begin{aligned} \mathbf{Q}(\underline{\theta}, \widehat{\underline{\theta}}^{(i)}) &= \mathbb{E} \left(-2\mathbf{L}(\underline{\theta}; X) \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right) \\ &= \sum_{t=1}^n \text{tr} \left\{ \text{diag}(R) \mathbb{E} \left[(Y_t - g(Y_{t-1})) (Y_t - g(Y_{t-1}))' \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right\} \\ &\quad + \sum_{t=1}^n \sum_{k=1}^K \left\{ -2 (\log \lambda_k) \mathbb{E} \left[Z_{t,k} \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + \mathbb{E} \left[Z_{t,k} Y_{t,k} \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right\} \\ &\quad + \sum_{t=1}^n \sum_{k=1}^K \varepsilon_t^2 \mathbb{E} \left[Z_{t,k} \exp(-Y_{t,k}) \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + n \sum_{k=1}^K \log Q_k, \\ &= \sum_{t=1}^n \text{tr} \left\{ \text{diag}(R) \left(\mathbb{E} \left[Y_t Y_t' \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] - \mathbb{E} \left[Y_t \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \alpha' - \mathbb{E} \left[Y_t \ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \beta' \right. \right. \\ &\quad \left. \left. - \alpha \mathbb{E} \left[Y_t' \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + \alpha \alpha' + \alpha \mathbb{E} \left[\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \beta' \right. \right. \\ &\quad \left. \left. - \beta \mathbb{E} \left[\left(\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \right) Y_t' \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + \beta \mathbb{E} \left[\left(\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \right) \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \alpha' \right. \right. \\ &\quad \left. \left. + \beta \mathbb{E} \left[\left(\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \right)^2 \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \beta' \right) \right\}, \\ &\quad + \sum_{t=1}^n \sum_{k=1}^K \left\{ -2 (\log \lambda_k) \mathbb{E} \left[Z_{t,k} \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + \mathbb{E} \left[Z_{t,k} Y_{t,k} \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right\} \\ &\quad + \sum_{t=1}^n \sum_{k=1}^K \varepsilon_t^2 \mathbb{E} \left[Z_{t,k} \exp(-Y_{t,k}) \mid \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + n \sum_{k=1}^K \log Q_k, \end{aligned} \tag{4.4}$$

Afin d'évaluer $\mathbf{Q}(\underline{\theta}, \widehat{\underline{\theta}}^{(i)})$, nous faisons appel aux algorithmes de filtrage et de lissage particulière de la même façon que pour les modèles *PAR-SV* univarié et multivarié.

Algorithme 4.1 (Algorithme de filtrage particulaire pour un modèle M -SV)

1. *Initialisation* : pour $j = 1, \dots, M$, générer $f_0^{(j)} \sim p_0(x_0)$ avec les poids initiaux $w_0^{(j)} = 1/M$, où M est le nombre de particules.

2. Pour $t = \overline{1, n}$,

(a) Pour $j = \overline{1, M}$,

i. Générer $e_t^{(j)} \sim \mathcal{N}(0, 1)$ et $\tilde{p}_t^{(j)} \sim \text{Mult}_K(1, \lambda)$.

ii. Calculer $p_t^{(j)} = \alpha + \beta \ln H_{t-1}^{(j)} + \sqrt{Q} e_t^{(j)}$.

iii. Mise à jour des poids : calculer

$$w_t^{(j)} = w_{t-1}^{(j)} p\left(\varepsilon_t | p_t^{(j)}, \tilde{p}_t^{(j)}\right) \propto w_{t-1}^{(j)} \sum_{k=1}^K \frac{1}{\sqrt{\exp p_{t,k}^{(j)}}} \exp\left(-\frac{\varepsilon_t^2}{2 \exp p_{t,k}^{(j)}}\right) \mathbf{1}_{\tilde{p}_{t,k}^{(j)}=1},$$

(b) Pour $j = \overline{1, M}$, normaliser les poids : calculer $\tilde{w}_t^{(j)} = w_t^{(j)} / \sum_{j=1}^M w_t^{(j)}$.

(c) Calculer la mesure de dégénérescence $n^{eff} = 1 / \sum_{j=1}^M \left(w_t^{(j)}\right)^2$,

i. si $n^{eff} \leq n^T$: $\left(f_t^{(j)}, \tilde{f}_t^{(j)}\right)$ par rééchantillonnage de $\left(p_t^{(1)}, \tilde{p}_t^{(1)}\right), \dots, \left(p_t^{(M)}, \tilde{p}_t^{(M)}\right)$ avec les poids $\tilde{w}_t^{(j)}$. Actualiser $\tilde{w}_t^{(j)} = 1/M$.

ii. sinon pour $j = \overline{1, M}$, $f_t^{(j)} = p_t^{(j)}$ et $\tilde{f}_t^{(j)} = \tilde{p}_t^{(j)}$.

(d) Pour $j = \overline{1, M}$, calculer $H_t^{(j)} = \sum_{k=1}^K \lambda_k \exp f_{t,k}^{(j)}$.

3. Finalement, la séquence des M particules $\left\{\left(f_t^{(j)}, \tilde{f}_t^{(j)}\right); j = \overline{1, M}\right\}$ est un échantillon de $(Y_t, Z_t | \mathcal{F}_t)$, pour $t = \overline{1, n}$.

Algorithme 4.2 (Algorithme de lissage particulaire pour un modèle M -SV)

1. Pour $j = \overline{1, M}$, choisir $\left(s_n^{(j)}, \tilde{s}_n^{(j)}\right) = \left(f_n^{(j)}, \tilde{f}_n^{(j)}\right)$, avec les probabilités $\tilde{w}_n^{(j)}$. Poser $W_n^{(j)} = 1/M$.

2. Pour $j = \overline{1, M}$,

(a) Pour $t = t-1, \dots, 0$, calculer pour $i = \overline{1, M}$

$$\begin{aligned} W_{t|t+1}^{(i)} &= \tilde{w}_t^{(i)} p\left(s_{t+1}^{(j)}, \tilde{s}_{t+1}^{(j)} \mid f_t^{(i)}, \tilde{f}_t^{(i)}\right) \\ &\propto \frac{\tilde{w}_t^{(i)}}{\prod_{k=1}^K Q_k} \exp\left[\frac{-1}{2} \left(s_{t+1}^{(j)} - g\left(f_t^{(i)}\right)\right)' \text{diag}(R) \left(s_{t+1}^{(j)} - g\left(f_t^{(i)}\right)\right)\right] \prod_{k=1}^K \lambda_k^{\tilde{s}_{t+1,k}^{(j)}}. \end{aligned}$$

(b) Pour $i = \overline{1, M}$, normaliser les poids modifiés : calculer $\widetilde{W}_{t-1|t}^{(i)} = W_{t-1|t}^{(i)} / \sum_{j=1}^M W_{t-1|t}^{(j)}$.

(c) Choisir $\left(s_{t-1}^{(j)}, \tilde{s}_{t-1}^{(j)}\right) = \left(f_{t-1}^{(i)}, \tilde{f}_{t-1}^{(i)}\right)$, avec les probabilités $\widetilde{W}_{t-1|t}^{(i)}$.

A la fin de l'étape de lissage particulière, nous avons $n \times M$ réalisations selon $Mult_K(1, \lambda)$. Ainsi, un estimateur pour λ_k est donné par

$$\widehat{\lambda}_k = \frac{\sum_{t=1}^n \sum_{j=1}^M \widetilde{s}_{t,k}^{(j)}}{n \times M} = \frac{\sum_{t=1}^n \widehat{Z}_{t,k}^{(n)}}{n}, \text{ pour tout } k = \overline{1, K-1},$$

qui converge presque sûrement selon le théorème ergodique.

Les dérivées premières partielles de (4.4), par rapport aux paramètres inconnus α , β et Q , sont données par

$$\begin{aligned} \frac{\partial \mathbf{Q}(\underline{\theta}, \widehat{\underline{\theta}}^{(i)})}{\partial \alpha} &= 2 \text{diag}(R) \sum_{t=1}^n \left(\alpha - \mathbb{E} \left[Y_t | \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + \beta \mathbb{E} \left[\left(\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \right) \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right), \\ \frac{\partial \mathbf{Q}(\underline{\theta}, \widehat{\underline{\theta}}^{(i)})}{\partial \beta} &= 2 \text{diag}(R) \sum_{t=1}^n \left\{ \alpha \mathbb{E} \left[\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right. \\ &\quad \left. - \mathbb{E} \left[\left(\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \right) Y_t \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] + \beta \mathbb{E} \left[\left(\ln \sum_{l=1}^K \lambda_l \exp(Y_{t-1,l}) \right)^2 \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right\}, \end{aligned}$$

et

$$\frac{\partial \mathbf{Q}(\underline{\theta}, \widehat{\underline{\theta}}^{(i)})}{\partial Q} = nR - \text{diag}(R^{\odot 2}) \sum_{t=1}^n \text{dig} \left(\mathbb{E} \left[(Y_t - g(Y_{t-1})) (Y_t - g(Y_{t-1}))' \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right).$$

L'estimateur $\widehat{\theta}^{(i+1)}$ du paramètre θ peut être obtenu en résolvant le système des équations

$$\partial \mathbf{Q}(\underline{\theta}, \widehat{\underline{\theta}}^{(i)}) / \partial \theta = 0.$$

Par conséquent, à la $(i+1)^{\text{ème}}$ itération, les estimateurs des paramètres de α , β et Q , sont respectivement donnés par

$$\begin{cases} \widehat{\beta} = \frac{\sum_{t=1}^n \mathbb{E} \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l \exp(Y_{t-1,l}) \right) Y_t \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] - \frac{1}{n} \left(\sum_{t=1}^n \mathbb{E} \left[Y_t | \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right) \sum_{t=1}^n \mathbb{E} \left[\ln \sum_{l=1}^K \widehat{\lambda}_l \exp(Y_{t-1,l}) \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right]}{\left(\sum_{t=1}^n \mathbb{E} \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l \exp(Y_{t-1,l}) \right)^2 \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] - \frac{1}{n} \left(\sum_{t=1}^n \mathbb{E} \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l \exp(Y_{t-1,l}) \right) \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right)^2 \right)}, \\ \widehat{\alpha} = \mathbb{E} \left[Y_t' \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] - \widehat{\beta} \sum_{t=1}^n \mathbb{E} \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l \exp(Y_{t-1,l}) \right) \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right], \\ \widehat{Q} = \frac{1}{n} \sum_{t=1}^n \text{dig} \left(\mathbb{E} \left[\left(Y_t - \widehat{\alpha} - \widehat{\beta} \ln \sum_{l=1}^K \widehat{\lambda}_l \exp(Y_{t-1,l}) \right) \left(Y_t - \widehat{\alpha} - \widehat{\beta} \ln \sum_{l=1}^K \widehat{\lambda}_l \exp(Y_{t-1,l}) \right)' \middle| \underline{\varepsilon}, \widehat{\underline{\theta}}^{(i)} \right] \right). \end{cases}$$

Ces estimateurs peuvent être évalués en approximant les espérances conditionnelles par des fonctions

des particules lissées comme suit

$$\left\{ \begin{array}{l} \widehat{\lambda}_k^{(i+1)} = \frac{1}{n \times M} \sum_{t=1}^n \sum_{j=1}^M \widetilde{s}_{t,k}^{(j)}, \\ \widehat{\beta}^{(i+1)} = \frac{\sum_{t=1}^n \sum_{j=1}^M \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l^{(i+1)} \exp(s_{t-1,l}^{(j)}) \right) s_t^{(j)} \right] - \frac{1}{n} \left(\sum_{t=1}^n \sum_{j=1}^M s_t^{(j)} \right) \sum_{t=1}^n \sum_{j=1}^M \left[\ln \sum_{l=1}^K \widehat{\lambda}_l^{(i+1)} \exp(Y_{t-1,l}) \right]}{\left(\sum_{t=1}^n \sum_{j=1}^M \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l^{(i+1)} \exp(s_{t-1,l}^{(j)}) \right) \right]^2 \right) - \frac{1}{n} \left(\sum_{t=1}^n \sum_{j=1}^M \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l^{(i+1)} \exp(s_{t-1,l}^{(j)}) \right) \right] \right)^2}, \\ \widehat{\alpha}^{(i+1)} = \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^M s_t^{(j)} - \widehat{\beta}^{(i+1)} \sum_{t=1}^n \sum_{j=1}^M \left[\left(\ln \sum_{l=1}^K \widehat{\lambda}_l^{(i+1)} \exp(s_{t-1,l}^{(j)}) \right) \right], \\ \widehat{Q}^{(i+1)} = \frac{1}{n} \sum_{t=1}^n \sum_{j=1}^M \text{diag} \left\{ \left(s_t^{(j)} + \widehat{P}_t^{(n)} - \widehat{\alpha}^{(i+1)} - \widehat{\beta}^{(i+1)} \ln \sum_{l=1}^K \widehat{\lambda}_l^{(i+1)} \exp(s_{t-1,l}^{(j)}) \right) \right. \\ \left. \left(s_t^{(j)} + \widehat{P}_t^{(n)} - \widehat{\alpha}^{(i+1)} - \widehat{\beta}^{(i+1)} \ln \sum_{l=1}^K \widehat{\lambda}_l^{(i+1)} \exp(s_{t-1,l}^{(j)}) \right)' \right\}, \end{array} \right.$$

où $\widehat{P}_t^{(n)} = \sum_{j=1}^M (s_t^{(j)} - \widehat{Y}_t^{(n)}) (s_t^{(j)} - \widehat{Y}_t^{(n)})' / M - 1$ et $\widehat{Y}_t^{(n)} = \sum_{j=1}^M s_t^{(j)} / M$.

Remarque 4.2 *La volatilité est inobservable mais peut être estimée en utilisant l'échantillon complet. En effet, elle peut être estimée par la volatilité lissée donnée par*

$$\widehat{h}_{t|n} = \frac{\sum_{k=1}^K \widehat{\lambda}_k \sum_{j=1}^M \exp s_{t,k}^{(j)}}{M}.$$

Elle peut être également estimée à partir des observations jusqu'à la période considérée (par la volatilité filtrée)

$$\widehat{h}_{t|t} = \frac{\sum_{k=1}^K \widehat{\lambda}_k \sum_{j=1}^M \widetilde{w}_t^{(j)} \exp f_{t,k}^{(j)}}{M}.$$

Il est important de noter que dans le cas des algorithmes particulières, nous avons choisi la loi d'évolution comme loi d'importance, i.e. $q(h_t | Y_{t-1}, \mathcal{F}_{t-1}) = f(h_t | Y_{t-1})$. En se basant sur ce choix, on peut approximer la variance conditionnelle $\text{Var}(\varepsilon_t | \mathcal{F}_{t-1})$ (voir Doucet et al. (2000) pour le cas général des modèles espace d'états non linéaire non gaussien, et Kim (2005) pour les modèles SV comme cas particulier des modèles espace d'états) par les mêmes estimations de h_t .

4.5.2 Exemple illustratif en utilisant des données simulées

Pour étudier la performance de la méthode proposée pour l'estimation des paramètres d'un modèle M -SV, nous avons simulé des données à partir de plusieurs modèles dont nous rapportons ici les résultats d'un seul modèle. Il s'agit d'un mélange de deux modèles SV. Notons que les paramètres du modèle simulé sont choisis de façon à ce que les paramètres vérifient la condition suffisante de stationnarité, i.e. $\max_{l=1, \dots, K} |\beta^{(l)}| < 1$. Notons également qu'afin d'examiner la convergence empirique des estimateurs, nous avons considéré différentes tailles ($n = 200, 500$ et 750). Pour chaque taille, nous avons simulé 200 séries. Les vraies valeurs (TV) des paramètres du processus générateur de données, la moyenne empirique (MLE) et l'écart-type empirique ($Sd-MLE$) de leurs estimations pour les 200 réplifications sont rapportées dans le Tableau 4.1.

Tableau 4.1. Résultats de simulation d'un mélange de modèles SV , avec $n = 200, 500$ et 750 .Le nombre de réplifications est égal à 200 et le nombre de particules $M = 100$.

	$n = 200$			$n = 500$		$n = 750$	
	TV	MLE	$Sd-MLE$	MLE	$Sd-MLE$	MLE	$Sd-MLE$
λ_1	0.25	0.2419	0.0599	0.2443	0.0314	0.2458	0.0305
α_1	1.0	1.1612	0.1399	1.1114	0.0854	1.0988	0.0600
α_2	0.4	0.2137	0.2546	0.2578	0.1716	0.2724	0.1387
β_1	0.9	0.7365	0.0842	0.7684	0.0547	0.7792	0.0461
β_2	-0.5	-0.2931	0.2221	-0.3458	0.1667	-0.3650	0.1242
Q_1	0.5	0.5665	0.0382	0.5666	0.0237	0.5651	0.0174
Q_2	0.8	0.8708	0.0842	0.8599	0.0513	0.8637	0.0411

À partir du Tableau 4.1, nous pouvons voir clairement que, même avec des séries de tailles relativement petites, le biais des estimations obtenues est généralement petit. Nous pouvons remarquer également que le biais diminue à chaque fois que la taille de la série augmente et les estimations sont de plus en plus proches des vraies valeurs. D'autre part, nous pouvons observer que les écart-types empiriques de certaines estimations de paramètres sont relativement importantes. Par exemple, lorsque la taille de la série est égale à 200, ce résultat est plus frappant pour le paramètre α_2 . Dans ce cas, la t -stat est d'environ 0.84. Cette statistique est augmentée à plus de 1.96 lorsque la taille de l'échantillon est augmentée à 750. D'une façon générale, à chaque fois que la taille de la série augmente, les biais ainsi que les écarts types empiriques diminuent, ce qui prouve la convergence empirique des estimateurs ML .

Conclusion générale et perspectives

Dans cette thèse, nous avons tenté une modeste contribution dans le domaine des séries chronologiques en proposant trois modèles à volatilité stochastique. Ces modèles permettent de bien décrire la volatilité instantanée des séries financières. Nous avons décomposé cette thèse en quatre chapitres. Dans le premier chapitre, nous avons présenté un bref aperçu sur les modèles espace d'états et sur deux méthodes d'estimation du vecteur d'état à savoir le filtre de Kalman et le filtrage-lissage particulière. Les deux chapitres suivants ont été consacrés à l'étude de deux modèles à volatilité stochastique (univarié et multivarié) à coefficients périodiques (*PAR-SV* et *PVAR-SV*). Le dernier chapitre présente une étude d'une nouvelle classe, à savoir la classe de mélange de modèles à volatilité stochastique (*M-SV*).

En somme, une analyse des propriétés probabilistes des modèles proposés a été établie et deux procédures d'estimation des paramètres basées sur les algorithmes exposés dans le premier chapitre, ont été élaborées. Des illustrations par des applications sur des données simulées et sur des données financières réelles ont été présentées. Une discussion des résultats de cette partie empirique confirme l'intérêt des modèles proposés et la bonne performance des méthodes d'estimation adoptées.

Bien que plusieurs résultats aient été obtenus dans cette thèse, il existe un certain nombre de zones inexploitées et quelques problèmes non résolus. Nous pouvons discuter des orientations futures communes à l'ensemble des modèles étudiés, en ouvrant des pistes d'analyse et proposer des horizons nouveaux, notamment,

1. Concernant les méthodes d'estimation, plusieurs travaux supplémentaires sont nécessaires. En comparaison avec les méthodes d'approximation standard (par exemple le filtre de Kalman étendu), le principal avantage des méthodes particulières est qu'elles ne reposent sur aucune technique de linéarisation locale ou d'approximation fonctionnelle brute. Le prix à payer pour cette flexibilité est le calcul. En effet, ces méthodes sont coûteuses en termes de temps. Effectuer une itération de la procédure d'estimation implique la génération de $2M$ échantillons, et le calcul de plusieurs quantités pour obtenir la vraisemblance attendue. En outre, une critique courante de l'algorithme *EM* est que sa convergence peut être assez lente.

Afin d'accélérer le processus, nous avons utilisé un nombre moindre de particules ($M = 100$, $M = 200$, pour l'étude de simulation). Malgré ces efforts, il faut relativement beaucoup de temps pour parvenir à la convergence. En plus, la vitesse de convergence et la précision sont deux objectifs en quelque sorte en conflit l'un avec l'autre (augmenter le nombre de particules, implique une bonne précision et le contraire est vrai). Nous comptons étudier comment accélérer la procédure d'estimation, en visant une meilleure précision, dans un futur proche. Dans ce cadre, nous pensons

à exploiter les champs suivants :

- (a) Appliquer certaines variantes du filtre particulaire qui ont été développées pour rendre le filtre particulaire plus robuste.
 - (b) Les filtres particulaires peuvent être considérés comme des alternatives puissantes aux algorithmes *MCMC*. Ils peuvent également être utilisés pour concevoir des schémas *MCMC* très efficaces.
 - (c) Exploiter les méthodes basées sur les filtres particulaires permettant l'utilisation des filtres de Kalman locaux (voir par exemple Chen et Liu, 2000 ; Frei et Künsch, 2013 ; Dahia et al., 2005). Le filtre *KPKF* (Kalman-Particle Kernel Filter), proposé par Dahia et al. (2004), combine à la fois les avantages du filtre de Kalman étendu (*EKF*) en termes de précision et du filtre particulaire régularisé (variante de filtre particulaire) en termes de robustesse. Ceci afin que les particules soient maintenues dans la zone d'intérêt de l'espace d'état. En utilisant cette technique, le nombre nécessaire pour le rééchantillonnage du système de particules est alors réduit. Il est à noter que le *KPKF* introduit un nouveau type de rééchantillonnage, d'une part pour préserver la structure du mélange, et d'autre part pour éviter la dégénérescence des poids des particules. Ces améliorations permettent au *KPKF* d'atténuer les fluctuations Monte Carlo.
 - (d) Augmenter le vecteur d'état en insérant les paramètres. L'état et les paramètres sont estimés simultanément par le filtrage-lissage particulaire (voir par exemple Kitagawa et Sato, 2001). Cette approche consiste à traiter les paramètres comme variables aléatoires admettant une densité de probabilité.
2. Dans les exemples de simulation, nous avons montré que les procédures d'estimation proposées fonctionnent bien et la performance d'estimation s'améliore significativement avec l'augmentation de la taille. Ce qui montre la propriété de consistance empirique. Ainsi, le comportement asymptotique des estimateurs proposé devrait être théoriquement étudié.
 3. Tout au long de cette thèse, nous avons supposé que toutes les données ont été enregistrées, une autre direction future importante consisterait à étudier l'estimation des paramètres de ces modèles lorsqu'une partie des données est manquante.
 4. Il pourrait être intéressant aussi d'étudier le modèle *M-SV* dans sa version multivariée. Et pourquoi pas étudier aussi le *M-SV* périodique, en combinant les avantages du modèle *SV* périodique et le mélange *SV* (comme fait dans les modèles type Mélange *ARCH*, voir par exemple Bentarzi et Hamdi, 2008a, b ; Hamdi et Souam 2013, 2017).
 5. Enfin, dans le cadre de cette thèse, notre intérêt s'est focalisé sur les modèles *SV* à temps discret. Cependant, Il serait intéressant d'étudier les modèles *SV* à temps continu.

Bibliographie

- [1] Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V. H., & Enriquez, I. (2010). Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions. *Computational statistics & data analysis*, **54**, 2883-2898.
- [2] Abramowitz, M. and Stegun, N. C., (1970). *Handbook of Mathematical Functions*, Dover Publications, New York.
- [3] Adams, G. J., & Goodwin, G. C. (1995). Parameter estimation for periodic *ARMA* models. *Journal of Time Series Analysis*, **16**, 127-145.
- [4] Aknouche, A., (2006). Etude théorique et algorithmique de la modélisation des séries chronologiques périodiques. Thèse de Doctorat. USTHB, Alger, Algérie.
- [5] Aknouche, A., (2007). Causality conditions and autocovariance calculations in *PVAR* models. *J. Statistical Computation and Simulation*, **77**, 769-780.
- [6] Aknouche, A. (2017). Periodic autoregressive stochastic volatility. *Statistical Inference for Stochastic Processes*, **20**, 139-177.
- [7] Aknouche A., Belbachir, H. & Hamdi, F., (2008). Note on calculating autocovariances of periodic *ARMA* models. *Communications in statistics-Simulation and computation*, **37**, 924-927.
- [8] Aknouche, A., & Bibi, A. (2009). Quasi-maximum likelihood estimation of periodic *GARCH* and periodic *ARMA-GARCH* processes. *Journal of Time Series Analysis*, **30**, 19-46.
- [9] Aknouche A., Bibi, A., & Hamdi, F. (2007). Periodic autoregressive stochastic volatility. *Journées de Statistique, Biskra*.
- [10] Aknouche, A. & Hamdi, F., (2007) Chandrasekhar-type recursions for periodic linear systems. *Far East Journal of Theoretical Statistics*, **22**, 65-80.
- [11] Aknouche, A. & Hamdi, F., (2008a). Extension du filtre de Chandrasekhar au cas des modèles espace d'états périodiques. *C. R. Acad. Sci. Ser. I*, **346**, 177-182.
- [12] Aknouche, A. & Hamdi, F., (2008b). Calculating the autocovariances and the likelihood calculations for periodic *VARMA* models, *J. Statistical Computation and simulation*.
- [13] Andel, J., (1983). Statistical analysis of periodic autoregression. *Aplikace Matimatiky*, **28**, 364-385.
- [14] Andersen, T. G. (1994). Stochastic autoregressive volatility : a framework for volatility modeling. *Mathematical finance*, **4**, 75-102.
- [15] Anderson, P. L. & Meerschaert, M. M., (2005). Parameter estimation for periodically stationary time series. *J. Time Ser. Anal.*, **26**, 489-518.

- [16] Anderson, B. D., & Moore, J. B. (1979). Optimal filtering. Englewood Cliffs, **21**, 22-95.
- [17] Ansley, C. F., (1979). An algorithm for the exact likelihood of a mixed autoregressive-moving average process. *Biometrika*, **66**, 59-65.
- [18] Andersen, T. G., & Sørensen, B. E. (1996). GMM estimation of a stochastic volatility model : a Monte Carlo study. *Journal of Business & Economic Statistics*, **14**, 328-352.
- [19] Aoki, M. (2013). State space modeling of time series. Springer Science & Business Media.
- [20] Arulampalam, M.S., Maskell, S., Gordon, N. & Clapp, T., (2002) . A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, **50**, 174-188.
- [21] Asai, M. (2009). Bayesian analysis of stochastic volatility models with mixture-of-normal distributions. *Mathematics and Computers in Simulation*, **79**, 2579-2596.
- [22] Asai, M., & McAleer, M. (2009a). The structure of dynamic correlations in multivariate stochastic volatility models. *Journal of Econometrics*, **150**, 182-192.
- [23] Asai, M., & McAleer, M. (2009b). Multivariate stochastic volatility, leverage and news impact surfaces. *The Econometrics Journal*, **12**, 292-309.
- [24] Asai, M., McAleer, M., & Yu, J. (2006). Multivariate stochastic volatility : a review. *Econometric Reviews*, **25**, 145-175.
- [25] Asai, M., & McAleer, M., (2011) .Alternative Asymmetric Stochastic Volatility Models. *Econometric Reviews*, **30**, 548-564.
- [26] Asai, M., McAleer, M., & Medeiros, M. C. (2012). Asymmetry and long memory in volatility modeling. *Journal of Financial Econometrics*, **10**, 495-512.
- [27] Ausín, M.C., & Galeano, P., (2007). Bayesian estimation of the Gaussian mixture *GARCH* model, *Comput. Statist. Data Anal.* **51**, 2636–2652
- [28] Azrak, R. & Mélard, G., (1998). The exact quasi-likelihood of time-dependent *ARMA* models. *J. Statistical Planning and Inference*, **68**, 31-46.
- [29] Basawa, I. V. & Lund, R., (2001). Large sample properties of parameters estimates for periodic *ARMA* models. *J. Time Ser. Anal.*, **22**, 1-13.
- [30] Bauwens, L., Hafner, C.M., & Rombouts, J.V.K., (2007). Multivariate mixed normal conditional heteroskedasticity. *Comput. Statist. Data Anal.* **51**, 3551–3566.
- [31] Bauwens, L., Laurent, S., & Rombouts, J. V., (2006a). Multivariate GARCH : A survey. *Journal of Applied Econometrics*, **21**, 79-109.
- [32] Bauwens, L., Rime, D. & Sucarrat, G. (2006b). Exchange rate volatility and the mixture of distribution hypothesis. *Empirical Economics*, **30**, 889-911.
- [33] Bentarzi, M., (1995). Modèles de séries chronologiques à coefficients périodiques. Thèse de Doctorat en sciences. USTHB, Alger, Algérie.
- [34] Bentarzi, M., (1998). Model-building problem of periodically m-variate moving average processes. *J. Multivariate Analysis*, **66**, 1-21.

- [35] Bentarzi, M. & Aknouche A. (2005). Calculation of the Fisher information matrix for periodic *ARMA* models. *Communications in Statistics-Theory and Methods*, **34**, 891-903.
- [36] Bentarzi, M. & Aknouche A., (2006). An on-line estimation algorithm for periodic autoregressive models. *Communications in Statistics-Theory and Methods*, **35**, 1495-1512.
- [37] Bentarzi, M. & Hallin, M. (1994). On the Invertibility of Periodic Moving Average Models. *J. Time Ser. Anal.*, **15**, 263-268.
- [38] Bentarzi, M. & Hallin, M. (1996). Locally optimal tests against periodic autoregression. *Econometric Theory*, **12**, 88-112.
- [39] Bentarzi, M. & Hallin, M. (1998). Spectral factorization of periodically correlated *MA*(1) processes. *J. Applied Probability*, **35**, 46-54.
- [40] Bentarzi, M. & Hamdi, F., (2008a). Mixture Periodic Autoregressive Conditional Heteroskedastic Models. *Comput. Statist. Data Anal.*, **53**, 1-16.
- [41] Bentarzi, M. & Hamdi, F., (2008b). Mixture Periodic Autoregression with Periodic *ARCH* Errors. *Advances and Applications in Statistics*, **8**, 219-246.
- [42] Bentarzi, M., Guerbyenne, H. & Hemis, R., (2008). Predictive density order selection of periodic *AR* models. *Communications in statistics-Simulation and computation*, **37**, 1167-1182.
- [43] Biernacki, C. (2009). Pourquoi les modèles de mélange pour la classification. *Revue de MODULAD*, **40**, 1-22.
- [44] Biernacki, C., Celeux, G., Echenim, A., Govaert, G., & Langrognet, F. (2006). Le logiciel MIXMOD d'analyse de mélange pour la classification et l'analyse discriminante. *La revue de Modulad*, **35**, 25-44.
- [45] Bittanti, S. Colaneri, P. & De Nicolao G. D. (1988). The difference periodic Riccati equation for the periodic prediction problem. *IEEE Trans. Autom. Control*, **33**, 706-712.
- [46] Bittanti, S. & De Nicolao, G. (1993). Spectral factorization of linear periodic systems with application to the optimal prediction of periodic *ARMA* models. *Automatica*, **29**, 517-522.
- [47] Bollerslev, T., (1986). Generalized autoregressive conditional heteroscedasticity. *J. Econometrics*. **51**, 307-327.
- [48] Bollerslev, T., & Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business & Economic Statistics*, **14**, 139-151.
- [49] Bojanczyk, A. W. Golub, G. & Van Dooren, P., (1992). The periodic Schur decomposition : Algorithms and applications, In F.T. Luk (ed.). *Proceedings SPIE Conference*, **1770**, 31-42.
- [50] Bordes, L., Mottelet, S., & Vandekerckhove, P. (2006). Semiparametric estimation of a two-component mixture model. *The Annals of Statistics*, **34**, 1204-1232.
- [51] Boteva, A. & Boshnakov, G.N., (1992). An algorithm for the computation of the theoretical autocovariances of a periodic autoregressive process, In Proceedings of the 8th Seminar on Statistical Data Analysis. Varna, 9-15.
- [52] **Boussaha**, N. & Hamdi, F. (2017). On periodic autoregressive stochastic volatility models, *Journal of Statistical Computation and Simulation*, DOI :10.1080/00949655.2017.1401626.

- [53] **Boussaha**, N., Hamdi, F. & SOUAM, S. (2017). Multivariate Periodic Stochastic Volatility Models : Applications to Algerian dinar exchange rates and oil prices modeling. Submitted.
- [54] **Boussaha**, N. & Hamdi, F. (2018). Mixture Stochastic Volatility Model. Mimeo
- [55] Brokwell, P.J. ; Davis, R.A., (1991). Time Series : Theory and Methods, (2nd edn), New York : Springer-Verlag.
- [56] Broto, C. & E. Ruiz (2004). Estimation Methods for Stochastic Volatility Models : A Survey, Journal of Economic Surveys. **18**, 613-649.
- [57] Boshnakov, G. N., (1996). Recursive computation of the parameters of periodic autoregressive moving-average processes. *J. Time Ser. Anal.*, **17**, 333-349.
- [58] Breidt, F. J., Crato, N., & De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. Journal of econometrics, **83**, 325-348.
- [59] Brockwell, P. J., & Davis, R. A. (2016). Introduction to time series and forecasting. springer.
- [60] Cappé, O., Godsill, S. J. & Moulines, E. (2007). An overview of existing methods and recent advances in sequential Monte Carlo. *IEEE Proceedings*, **95**,899-924.
- [61] Carnero, M. A., Peña, D., & Ruiz, E. (2004). Persistence and kurtosis in GARCH and stochastic volatility models. *Journal of Financial Econometrics*, **2**, 319-342.
- [62] Carvalho, C. M., & Lopes, H. F. (2007). Simulation-based sequential analysis of Markov switching stochastic volatility models. *Computational Statistics & Data Analysis*, **51**, 4526-4542.
- [63] Casals, J., Sotoca, S. & Jerez, M., (1998). Un algoritmo rápido para evaluar la verosimilitud exacta de modelos *VARMAX* periódicos, *Estadística Espanola*, **40**, 269-291.
- [64] Casella, G., & Robert, C. P. (1996). Rao-Blackwellisation of sampling schemes. *Biometrika*, **83**, 81-94.
- [65] Cavanaugh, J. & Shumway, R. H., (1997). A bootstrap-corrected variant of aic for state-space model selection. *Statistica Sinica*, **7**, 473-496.
- [66] Cavanaugh, J. E., & Shumway, R. H. (1996). On computing the expected Fisher information matrix for state-space model parameters. *Statistics & probability letters*, **26**, 347-355.
- [67] Chan, D., Kohn, R., & Kirby, C. (2006). Multivariate stochastic volatility models with correlated errors. *Econometric Reviews*, **25**, 245-274.
- [68] Chen, R., & Liu, J. S. (2000). Mixture kalman filters. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, **62**, 493-508.
- [69] Chen, C. W., Liu, F. C., & So, M. K. (2008). Heavy-Tailed-Distributed Threshold Stochastic Volatility Models In Financial Time Series. *Australian & New Zealand Journal of Statistics*, **50**, 29-51.
- [70] Chib, S., Nardari, F., & Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, **108**, 281-316.
- [71] Chib, S., Nardari, F., & Shephard, N. (2006). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* ,**134**, 341-371.

- [72] Cleveland, W. P & Tiao, G. C. (1979). Modeling seasonal time series. *Revue Economique Appliquée*, **32**, 107-129.
- [73] Dahia, K. (2005). Nouvelles méthodes en filtrage particulaire-Application au recalage de navigation inertielle par mesures altimétriques (Doctoral dissertation, Université Joseph-Fourier-Grenoble I).
- [74] Dahia, K., Musso, C., Pham, D. T., & Guibert, J. P. (2004, September). Application of the Kalman-Particle Kernel Filter to the updated inertial navigation system. In Signal Processing Conference, 2004 12th European (pp. 601-604). *IEEE*.
- [75] Dahlhaus, R. & Subba Rao, S., (2006). Statistical inference for time-varying *ARCH* processes. *J. Ann. Statist.*, **34** , 1075-1114.
- [76] Danielsson, J. (1998). Multivariate stochastic volatility models : estimation and a comparison with *VGARCH* models. *Journal of Empirical Finance*, **5**, 155-173.
- [77] Davis, R. A. & Yam, G. R. (2005) Estimation for state space models based on a likelihood approximation. *Statistica Sinica*, **15**, 381-406.
- [78] Day, N.E. (1969). Estimating the components of a mixture of two normal distributions. *Biometrika* **56**, 463-474.
- [79] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the *EM* algorithm. *Journal of the royal statistical society. Series B (methodological)* **39**, 1-38.
- [80] Ding, L., & Vo, M. (2012). Exchange rates and oil prices : A multivariate stochastic volatility analysis. *The Quarterly Review of Economics and Finance*, **52**, 15-37.
- [81] Doucet, A. Godsill, S. & Andrieu, C., (2000). On Sequential Monte Carlo Sampling Methods for Bayesian Filtering. *Statistics and Computing*, **10**, 197–208.
- [82] Doucet, A., De Freitas, N., & Gordon, N. (2001). Sequential Monte Carlo methods in practice. Springer, New York, NY.
- [83] Doucet, A. Johansen. Adam M., (2012). A Tutorial on Particle Filtering and Smoothing Fifteen years later. in Handbook of Nonlinear Filtering (eds. D. Crisan et B. Rozovsky), Oxford University Press.
- [84] Durbin, J. & Koopman, S. J., (2001). *Time Series Analysis by State Space Methods*. Oxford, New York.
- [85] Efron, B., (1979). Bootstrap methods : Another look at the jackknife. *Annals of Statistics*, **7**, 1-26.
- [86] Elliott, R. J., Liew, C. C., & Siu, T. K. (2011). On filtering and estimation of a threshold stochastic volatility model. *Applied Mathematics and Computation*, **218**, 61-75.
- [87] Engle, R., (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of the United Kingdom inflation. *Econometrica*, **50**, 987-1007.
- [88] Fama E., (1965) The behaviour of stock market prices. *Journal of Business* **38**, 34–105.
- [89] Fearnhead, P. (2008). Computational methods for complex stochastic systems : A review of some alternatives to MCMC. *Statistics and Computing*, **18**, 151-171.

- [90] Figueiredo, M. A. T., & Jain, A. K. (2002). Unsupervised learning of finite mixture models. *IEEE Transactions on pattern analysis and machine intelligence*, **24**, 381-396.
- [91] Fong, P. W., Li, W. K., Yau, C. W. & Wong, C. S., (2007). On a mixture vector autoregressive model. *Canadian Journal of Statistics*, **35**, 135-150.
- [92] Franses, P. H. (1996). Periodicity and Stochastic Trends in Economic Times Series. Oxford University Press.
- [93] Franses, P. H., & Paap, R. (2000). Modelling day-of-the-week seasonality in the S&P 500 index. *Applied Financial Economics*, **10**, 483-488.
- [94] Franses, P. & Paap, R. (2004). Periodic time series models. Oxford University Press.
- [95] Fridman, M., & Harris, L. (1998). A maximum likelihood approach for non-Gaussian stochastic volatility models. *Journal of Business & Economic Statistics*, **16**, 284-291.
- [96] Frei, M., & Künsch, H. R. (2013). Bridging the ensemble Kalman and particle filters. *Biometrika*, **100**, 781-800.
- [97] Freedman, D. A., (1981). Bootstrapping regression models. *Annals of Statistics*, **9**, 1218-1228.
- [98] Freedman, D. A. & Peters, S. C., (1984a). Bootstrapping a regression equation : Some empirical results. *J. American Statistical Association*, **79**, 97-106.
- [99] Freedman, D. A. & Peters, S. C., (1984b). Bootstrapping an econometric model : Some empirical results. *J. Business and Economic Statistics*, **2**, 150-158.
- [100] Gautier, A., (2005). Influence asymptotique de la correction par la moyenne sur l'estimation d'un modèle $AR(1)$ périodique, *C. R. Acad. Sci. Paris- Mathématique*, **340**, 315-318.
- [101] Gardner, G., Harvey, A. C. & Phillips, G. D. A. (1980). Algorithm AS 154. An algorithm for exact maximum likelihood estimation of autoregressive-moving average models by means of Kalman filtering. *Applied Statistics*, **29**, 311-322.
- [102] Ghosh, H., Gurung, B., & Prajneshu. (2015). Kalman filter-based modelling and forecasting of stochastic volatility with threshold. *Journal of Applied Statistics*, **42**, 492-507.
- [103] Giannikis, D., Vrontos, I. D. & Dellaportas, P., (2008). Modelling nonlinearities and heavy tails via threshold normal mixture *GARCH* models. *Comput. Statist. Data Anal.* **52**, 1549-1571.
- [104] Gladyshev, E. G. (1961). Periodically correlated random sequences. *Soviet. Math.*, **2**, 385-88.
- [105] Godsill, S., Doucet, A., & West, M. (2004). Monte Carlo smoothing for non-linear time series. *Journal of the American Statistical Association*, **199**, 156-168.
- [106] Gordon, N. J., Salmond, D. J., & Smith, A. F. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F (Radar and Signal Processing)*, **140**, 107-113.
- [107] Guerbyenne, H. (2007). Modèles de Séries Chronologiques Conditionnellement Hétéroscédastiques. Thèse de Doctorat d'Etat. U.S.T.H.B, Alger, Algérie.
- [108] Guerbyenne, H., & Hamdi, F. (2015). Bootstrapping Periodic State-Space Models. *Communications in Statistics-Simulation and Computation*, **44**, 374-401.

- [109] Gustafsson, F. (2010). Particle Filter Theory and Practice with Positioning Applications. *IEEE Aerospace and Electronic Systems Magazine*, 25, 7, 53-81.
- [110] Ghysels, E., Harvey, A. C., & Renault, E. (1996). 5 Stochastic volatility. *Handbook of statistics*, 14, 119-191.
- [111] Haas, M. & Paoletta, M.S. (2012). Mixture and Regime-Switching *GARCH* Models. *Handbook of volatility models and their applications*, 71-102.
- [112] Hamdi, F. (2008). Modèles espace d'états et processus périodiquement corrélés. Thèse de Doctorat en sciences. USTHB, Alger, Algérie.
- [113] Hamdi, F., & Souam, S. (2013). Mixture periodic *GARCH* models : Applications to exchange rate modeling. In *Modeling, Simulation and Applied Optimization (ICMSAO)*, 2013 5th International Conference on (pp. 1-6). *IEEE*.
- [114] Hamdi, F., & Souam, S. (2017). Mixture periodic *GARCH* models : theory and applications. *Empirical Economics*, 1-32.
- [115] Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57, 357-384.
- [116] Hamilton, J. D., (1994). *Time series analysis*. Princeton University Press, New Jersey.
- [117] Harvey, A. C., (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, New York.
- [118] Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, 14, 429-434.
- [119] Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, 61, 247-264.
- [120] Hensch, J. J. & Laub, A. J., (1994). Numerical solution of the discrete-time periodic Riccati equation, *IEEE Trans. Autom. Control*, 39, 1197-1210.
- [121] Hurvich, C. M., Shumway, R. H. & Tsai, C. L. (1990). Improved estimators of Kullback-Leibler information for autoregressive model selection in small samples. *Biometrika*, 77, 709-719.
- [122] Hurvich, C. M. & Tsai, C. L., (1993). A corrected Akaike information criterion for vector autoregressive model selection. *J. Time Ser. Anal.*, 14, 271-279.
- [123] Hürzeler, M., & Künsch, H. R. (1998). Monte Carlo Approximations for General State Space Models. *Journal of Computational and Graphical Statistics*, 7, 175-193.
- [124] Jacquier, E., Polson, N. G., & Rossi, P. E. (2002). Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics*, 20, 69-87.
- [125] Jimenez, C., McLeod, A.I. & Hipel, K.W., (1989) Kalman filter estimation for periodic autoregressive-moving average models. *Stochastic Hydrology and Hydraulics*, 3, 227-240.
- [126] Jungbacker, B., & Koopman, S. J. (2006). Monte Carlo likelihood estimation for three multivariate stochastic volatility models. *Econometric Reviews*, 25, 385-408.
- [127] Kalman E., (1960). A new approach to linear filtering and predicting problems. *Transactions of the ASME, Journal of Basic Engineering Series*, 82, 35-45.

- [128] Kim, J. (2005). Parameter Estimation in Stochastic Volatility Models with Missing Data Using Particle Methods and the EM Algorithm. PhD thesis. University of Pittsburgh, Pittsburgh.
- [129] Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility : likelihood inference and comparison with ARCH models. *The review of economic studies*, **65**, 361-393.
- [130] Kim, J., & Stoffer, D. S. (2008). Fitting stochastic volatility models in the presence of irregular sampling via particle methods and the em algorithm. *Journal of time series analysis*, **29**, 811-833.
- [131] Kitagawa, G. (1996). Monte Carlo Filter and Smoother for Nonlinear Non-Gaussian State-Space Models. *Journal of Computational and Graphical Statistics*, **5**, 1-25.
- [132] Kitagawa, G., & Sato, S. (2001). Monte Carlo smoothing and self-organising state-space model. In Sequential Monte Carlo methods in practice (pp. 177-195). Springer, New York, NY.
- [133] Klein, A. & Mélard, G. (1990). Fisher's information matrix for seasonal autoregressive moving average models. *J. Time Ser. Anal.*, **11**, 231-237.
- [134] Klein, A., Mélard, G., & Zahaf, T. (2000). Construction of the exact Fisher information matrix of Gaussian time series models by means of matrix differential rules. *Linear Algebra and its Applications*, **321**, 209-232.
- [135] Kollo, T. (2008). Multivariate skewness and Kurtosis measures with an application in ICA. *Journal of Multivariate Analysis*, **99**, 2328-2338.
- [136] Kollo, T., & von Rosen, D. (2006). *Advanced multivariate statistics with matrices* (Vol. **579**). Springer Science & Business Media.
- [137] Kong, A. (1992). A note on importance sampling using standardized weights. University of Chicago, Dept. of Statistics, Tech. Rep, 348.
- [138] Kong, A., Liu, J. S., & Wong, W. H. (1994). Sequential imputations and Bayesian missing data problems. *Journal of the American statistical association*, **89**, 278-288.
- [139] Künsch, H. R. (2001). State-space and hidden Markov models. in Complex Stochastic Systems (eds. O. E. Barndorff-Nielsen, D. R. Cox and C. Kluppelberg), CRC Press, 109-173.
- [140] Künsch, H. R. (2013). Particle filters. *Bernoulli*, **19**, 1391-1403.
- [141] Le, N.D., Martin, R.D., Raftery, A.E., (1996). Modeling flat stretches, bursts, and outliers in time series using mixture transition distribution models. *J. Amer. Statist. Assoc.*, **91**, 1504-1514.
- [142] Leisch, F. (2004). Flexmix : A general framework for finite mixture models and latent glass regression in R.
- [143] Li, W. K & Hui, Y. V. (1988). An Algorithm for the exact likelihood of periodic autoregressive moving average models. *Communication in Statistics- Simulation and Computation*, **16**, 1483-1494.
- [144] Liu, J. S. (1996). Metropolized independent sampling with comparisons to rejection sampling and importance sampling. *Statistics and Computing*, **6**, 113-119.
- [145] Ljung, L. & Caines, P. E. (1979). Asymptotic normality of prediction error estimators for approximate system models. *Stochastics*, **3**, 29-46.

- [146] Louis, T., (1982). Finding the observed information matrix when using the *EM* algorithm. *J. Roy. Statist. Soc. B*, **44**, 226–233.
- [147] Lund, R. & Basawa, I. V., (1999). Modeling and inference for periodically correlated time series. In : Gosh, S. (Ed.), *Asymptotics, Nonparametrics and Time series*, Marcel Dekker, New York, 37-62.
- [148] Lund, R. & Basawa, I. V. (2000). Recursive prediction and likelihood evaluation for periodic *ARMA* models. *J. Time Ser. Anal.*, **21**, 75-93.
- [149] Lütkepohl, H., (1991). *Introduction to multiple time series analysis*, Berlin : Springer Verlag.
- [150] Lütkepohl, H., (1996). *Handbook of matrices*. John Wiley&Sons.
- [151] Mandelbrot B., (1963). The variation of certain speculative prices. *Journal of Business* **36**, 394-419.
- [152] Mao, X., Ruiz, E., & Veiga, H. (2017). Threshold stochastic volatility : Properties and forecasting. *International Journal of Forecasting*, **33**, 1105-1123.
- [153] Marron, J. S., & Wand, M. P. (1992). Exact mean integrated squared error. *The Annals of Statistics*, 712-736.
- [154] McLachlan, G.J. & Krishnam, T., (1997). *The EM algorithm and extensions*. Wiley, New York.
- [155] McLachlan, G., & Peel, D. (2004). *Finite mixture models*. *John Wiley & Sons*.
- [156] Mélard G., (1984). Algorithm AS197 : A fast algorithm for the exact likelihood of autoregressive-moving average models. *Appl. Statistics*, **33**, 104-114.
- [157] Mélard G. & Klein A., (1994). On a fast algorithm for the exact information matrix of a Gaussian *ARMA* time series. *IEEE Trans. Signal Process*, **42**, 2201-2203.
- [158] Morf M., Sidhu G. S. & Kailath T., (1974). Some new algorithms for recursive estimation in constant, linear, discrete-time systems, *IEEE Trans. Autom. Control*, **19**, 315-323.
- [159] Nakamori, S., Hermoso, A. Jimknez, J. & Linares, J.,(2004). Chandrasekhar-type filter for a wide-sense stationary signal from uncertain observations using covariance information. *Appl. Math. Comput.*, **151**, 315-325.
- [160] Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns : A new approach. *Econometrica : Journal of the Econometric Society*, 347-370.
- [161] Newton, H. J. (1982). Using periodic autoregressions for multiple spectral estimation. *Technometrics*, **6**, 1310-1317.
- [162] Omori, Y., Chib, S., Shephard, N., & Nakajima, J. (2007). Stochastic volatility with leverage : Fast and efficient likelihood inference. *Journal of Econometrics*, **140**, 425-449.
- [163] Osborn, D. R. (1992). The implication of periodically varying coefficients for seasonal time series processes. *J. Econometrics*, **48**, 373-384.
- [164] Pagano, M., (1978). On periodic and multiple autoregression. *Annals of Statistics*, **6**, 1310-1317.
- [165] Pham, D. T. (2001). Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly weather review*, **129**, 1194-1207.

- [166] Pitt, M.K. & Shephard, N., (1999). Filtering via simulation : Auxiliary particle filters. *Journal of the American Statistical Association*, **94**, 590-599.
- [167] Pederzoli, C. (2006). Stochastic volatility and *GARCH* : a comparison based on UK stock data. *European Journal of Finance*, **1**, 41-59.
- [168] Pearlman, J. G. (1980). An algorithm for the exact likelihood of a high-order autoregressive moving average process. *Biometrika*, **67**, 232-233.
- [169] Platanioti, K. McCoy, E. J. & Stephens, D. A., (2005) . A Review of Stochastic Volatility : univariate and multivariate models. Imperial College London.
- [170] Pulgarin, L. (2001). Applied comparison between *GARCH* and stochastic volatility models. Master's thesis, University of Pittsburgh.
- [171] Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of econometrics*, **63**, 289-306.
- [172] Salas, J. D, Duane C. Boes & Smith, R. A. (1983). Estimation of *ARMA* Models with Seasonal Parameters. *Water Resources Research*, **18**, 1006-1010.
- [173] Sayed A. H. & Kailath T., (1994). Extended Chandrasekhar recursions. *IEEE Trans. Autom. Control*, **39**, 619-623.
- [174] Sayed, A. H., Kailath, T. Lev-Ari, H., (1994). Generalized Chandrasekhar recursions from the generalized Schur algorithm, *IEEE Trans. Autom. Control*, **39**, 2265-2269.
- [175] Shao Q., (2006). Mixture periodic autoregressive time series models. *Statist. Probab. Lett.*, **76**, 609-618.
- [176] Shao, Q. & Lund, R. (2004). Computation and characterization of autocorrelations and partial autocorrelations in periodic *ARMA* models. *J. Time Ser. Anal.*, **25**, 359-372.
- [177] Shea, B. L. (1984). Maximum likelihood estimation of multivariate *ARMA* process via the Kalman filter. *J. Time Ser. Anal.*, 91-101.
- [178] Shea, B. L. (1987). Estimation of multivariate time series. *J. Time Ser. Anal.*, **8**, 95-109.
- [179] Shea, B. L. (1989). The exact likelihood of a vector autoregressive moving average model. *J. Royal Statistical Society C*, **38**, 161-184.
- [180] Shephard, N. (1996). Statistical aspects of *ARCH* and stochastic volatility. In Cox, Hinkley, and Barndorff-Nielson, editors, *Time Series Models : In econometrics, finance and other fields*. Chapman and Hall, London.
- [181] Schweppe, F. (1965). Evaluation of likelihood function for Gaussian signals. *IEEE Trans. Inform. Theory*, **11**, 61-70.
- [182] Shumway, R. H., & Stoffer, D. S. (2000). Time series analysis and its applications. *Studies In Informatics And Control*, **9**, 375-376.
- [183] Smith, M., & Pitts, A. (2006). Foreign exchange intervention by the Bank of Japan : Bayesian analysis using a bivariate stochastic volatility model. *Econometric Reviews*, **25**, 425-451.
- [184] So, M. E. P., Lam, K., & Li, W. K. (1998). A stochastic volatility model with Markov switching. *Journal of Business & Economic Statistics*, **16**, 244-253.

- [185] Sreedhar, J. & Van Dooren, P., (1993). Periodic Schur form and some matrix equations. In U. Helmke, R. Mennicken and J. Saurer (eds), Proc. MTNS'93, Regensburg, Germany, **I**, 339–362.
- [186] Stoffer D. S. & Wall, K. D., (1991). Bootstrapping state-space models : Gaussian maximum likelihood estimation and the Kalman filter. *J. Amer. Statis Assoc.*, **86**, 1024-33.
- [187] Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. *Time series analysis : theory and practice*, **1**, 203-226.
- [188] Taylor, S. J. (1986), Modelling Financial Time Series. Chichester, New York, Wiley.
- [189] Tiao, G. C & Grupe, M. R., (1980). Hidden periodic autoregressive-moving average models in time series data. *Biometrika*, **67**, 365-373.
- [190] Taylor, S. J. (1994). Modeling stochastic volatility : A review and comparative study. *Mathematical finance*, **4**, 183-204.
- [191] Titterington, D.M., Smith, A.F.M., & Makov, U.E. (1985). Statistical Analysis of Finite Mixture Distributions. New York : Wiley.
- [192] Tong, H. (1978). On a threshold model. In C.H. Chen (ed.) , Pattern Recognition and Signal Processing. Amsterdam : SijthofJ and NoordhofJ, 575-586.
- [193] Tong, H., (1990). Nonlinear Time Series : a Dynamical System Approach. Oxford University Press, Oxford.
- [194] Tsiakas, I. (2006). Periodic Stochastic Volatility and Fat Tails. *Journal of Financial Econometrics*, **4**, 90-135.
- [195] Tukey, J.W. (1960). A survey of sampling from contaminated distributions. In Contributions to Probability and Statistics, I. Olkin, S.G. Ghurye, W.Hoeffding, W.G.
- [196] Ula, T. A. & Smadi A. A., (1997). Periodic stationary conditions for periodic autoregressive moving average processes as eigenvalue problems. *Water Resources Research*, **33**, 1929-1934.
- [197] Varga, A., (1997) Periodic Lyapunov equations : some applications and new algorithms. *Int. J. Control*, **67**, 69-87.
- [198] Varga, A. & Van Dooren, P., (2001). Computational methods for periodic systems - an overview. Proc. of IFAC Workshop on Periodic Control Systems, Como, Italy, 171-176.
- [199] Vecchia, A. V. (1985a). Maximum likelihood estimation for periodic autoregressive Moving average models. *Technometrics*, **27**, 375-384.
- [200] Vecchia, A. V. (1985b). Periodic autoregressive-moving average (*PARMA*) modeling with application to water resources. *Water Resources Bulletin*, **21**, 721-730.
- [201] Wall, K. D. & Stoffer D. S., (2002). A state space approach to bootstrapping conditional forecasts in *ARMA* models. *J. Time Ser. Anal.*, **23**, 733-751.
- [202] Wiggins, J. B. (1987). Option values under stochastic volatility : Theory and empirical estimates. *Journal of financial economics*, **19**, 351-372.
- [203] Wong, C. & Li,W., (2000). On a mixture autoregressive model. *J. Roy. Statist. Soc. Ser. B*, **62**, 95-115.

- [204] Wong, C. & Li,W., (2001). On a mixture autoregressive conditional heteroscedastic model. *J. Amer. Statist. Assoc.*, **96**, 982-995.
- [205] Xu, D. (2007), Asset Returns, Volatility and Value-at-Risk, Ph.D Thesis, The University of Western Ontario.
- [206] Xu, D., & Knight, J. (2013). Stochastic volatility model under a discrete mixture-of-normal specification. *Journal of Economics and Finance*, **37**, 216-239.
- [207] Yano, K. (2008). A self-organizing state space model and simplex initial distribution search. *Computational Statistics*, **23**, 197-216.
- [208] Yao, J. F. (2000). On least squares estimation for stable nonlinear *AR* processes. *Annals of the Institute of Statistical Mathematics*, **52**, 316-331.
- [209] Yao, J. F., & Attali, J. G. (2000). On stability of nonlinear *AR* processes with Markov switching. *Advances in Applied Probability*, **32**, 394-407.
- [210] Yu, J. (2012). Simulation-based estimation methods for financial time series models. In Handbook of Computational Finance pp, 401-435. Springer, Berlin, Heidelberg.
- [211] Yu, J., & Meyer, R., (2006) . Multivariate Stochastic Volatility Models : Bayesian Estimation and Model Comparison. *Econometric Reviews*, **25**, 361-384.
- [212] Zaritskii, V.S., Svetnik, V.B. & Šimelevič L.I., (1975) . Monte Carlo technique in problems of optimal data processing. *Automation and Remote Control*, **12**, 95-103.
- [213] Zhang, Z., Li,W.K., & Yuen, K.C., (2006). On a mixture *GARCH* time-series model. *J. Time Ser. Anal.* **27**, 577–597.