

Université de Montréal

**Modélisation des données financières par les modèles à chaîne de Markov
cachée de haute dimension**

par

Kassimou Abdoul Haki Maoude

Département de mathématiques et de statistique

Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures et postdoctorales
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en Statistique

avril 2022

Université de Montréal

Faculté des études supérieures et postdoctorales

Cette thèse intitulée

**Modélisation des données financières par les modèles à chaîne de Markov
cachée de haute dimension**

présentée par

Kassimou Abdoul Haki Maoude

a été évaluée par un jury composé des personnes suivantes :

Manuel Morales

(président-rapporteur)

Maciej Augustyniak

(directeur de recherche)

Mylène Bédard

(co-directeur)

Philippe Gagnon

(membre du jury)

Yang Lu

(examineur externe)

Fabian Bastin

(représentant du doyen de la FESP)

Thèse acceptée le :

08 Avril 2022

Sommaire

La classe des modèles à chaîne de Markov cachée (HMM, *Hidden Markov Models*) permet, entre autres, de modéliser des données financières. Par exemple, dans ce type de modèle, la distribution du rendement sur un actif financier est exprimée en fonction d'une variable non-observée, une chaîne de Markov, qui représente la volatilité de l'actif. Notons que les dynamiques de cette volatilité sont difficiles à reproduire, car la volatilité est très persistante dans le temps. Les HMM ont la particularité de permettre une variation de la volatilité selon les états de la chaîne de Markov. Historiquement, ces modèles ont été estimés avec un nombre faible de régimes (états), car le nombre de paramètres à estimer explose rapidement avec le nombre de régimes et l'optimisation devient vite difficile. Pour résoudre ce problème une nouvelle sous-classe de modèles à chaîne de Markov cachée, dite à haute dimension, a vu le jour grâce aux modèles dits factoriels et à de nouvelles méthodes de paramétrisation de la matrice de transition. L'objectif de cette thèse est d'étendre cette classe de modèles avec de nouvelles approches plus générales et de montrer leurs applications dans le domaine financier.

Dans sa première partie, cette thèse formalise la classe des modèles factoriels à chaîne de Markov cachée et étudie les propriétés théoriques de cette classe de modèles. Dans ces modèles, la dynamique de la volatilité dépend d'une chaîne de Markov latente de haute dimension qui est construite en multipliant des chaînes de Markov de dimension plus faible, appelées composantes. Cette classe englobe les modèles factoriels à chaîne de Markov cachée précédemment proposés dont les composantes sont de dimension deux. Le modèle MDSV (*Multifractal Discrete Stochastic Volatility*) est introduit afin de pouvoir considérer des composantes de dimension supérieure à deux, généralisant ainsi les modèles factoriels existants. La paramétrisation particulière de ce modèle lui offre suffisamment de flexibilité pour reproduire différentes allures de décroissance de la fonction d'autocorrélation, comme celles qui sont observées sur les données financières. Un cadre est également proposé pour modéliser séparément ou simultanément les données de rendements financiers et de variances réalisées. Une analyse empirique sur 31 séries d'indices financiers montre que le modèle MDSV présente de meilleures performances en termes d'estimation et de prévision par rapport au modèle *realized EGARCH*.

La modélisation par l'entremise des modèles factoriels à chaîne de Markov cachée nécessite qu'on définisse le nombre N de composantes à multiplier et cela suppose qu'il n'existe pas d'incertitude lié à ce nombre. La seconde partie de cette thèse propose, à travers une approche bayésienne, le modèle iFHMV (*infinite Factorial*

Hidden Markov Volatility) qui autorise les données à déterminer le nombre de composantes nécessaires à leur modélisation. En s'inspirant du processus du buffet indien (IBP, *Indian Buffet Process*), un algorithme est proposé pour estimer ce modèle, sur les données de rendements financiers. Une analyse empirique sur les données de deux indices financiers et de deux actions permet de remarquer que le modèle iFHMV intègre l'incertitude liée au nombre de composantes pour les estimations et les prévisions. Cela lui permet de produire de meilleures prévisions par rapport à des modèles de référence.

Mots clés : Volatilité, modèles à chaîne de Markov cachée, rendements financiers, variances réalisées, persistance, modèles joints, modèles bayésiens, processus du buffet indien, construction rupture de bâton.

Summary

Hidden Markov Models (HMMs) are popular tools to interpret, model and forecast financial data. In these models, the return dynamics on a financial asset evolve according to a non-observed variable, a Markov chain, which generally represents the volatility of the asset. This volatility is notoriously difficult to reproduce with statistical models as it is very persistent in time. HMMs allow the volatility to vary according to the states of a Markov chain. Historically, these models are estimated with a very small number of regimes (states), because the number of parameters to be estimated grows quickly with the number of regimes and the optimization becomes difficult. The objective of this thesis is to propose a general framework to construct HMMs with a richer state space and a higher level of volatility persistence.

In the first part, this thesis studies a general class of high-dimensional HMMs, called factorial HMMs, and derives its theoretical properties. In these models, the volatility is linked to a high-dimensional Markov chain built by multiplying lower-dimensional Markov chains, called components. We discuss how previously proposed models based on two-dimensional components adhere to the factorial HMM framework. Furthermore, we propose a new process—the Multifractal Discrete Stochastic Volatility (MDSV) process—which generalizes existing factorial HMMs to dimensions larger than two. The particular parametrization of the MDSV model allows for enough flexibility to reproduce different decay rates of the autocorrelation function, akin to those observed on financial data. A framework is also proposed to model financial log-returns and realized variances, either separately or jointly. An empirical analysis on 31 financial indices reveals that the MDSV model outperforms the realized EGARCH model in terms of fitting and forecasting performance.

Our MDSV model requires us to pre-specify the number of components and assumes that there is no uncertainty on that number. In the second part of the thesis, we propose the infinite Factorial Hidden Markov Volatility (iFHMV) model as part of a Bayesian framework to let the data drive the selection of the number of components and take into account the uncertainty related to the number of components in the fitting and forecasting procedure. We also develop an algorithm inspired by the Indian Buffet Process (IBP) to estimate the iFHMV model on financial log-returns. Empirical analyses on two financial indices and two stocks show that the iFHMV model outperforms popular benchmarks in terms of forecasting performance.

Keywords: Volatility, hidden Markov models, financial return, realized variance, persistence, joint models, bayesian models, Indian buffet process, stick-breaking construction.

Table des matières

| | |
|---|-------|
| Sommaire | v |
| Summary | vii |
| Liste des tableaux | xv |
| Table des figures | xvii |
| Liste des sigles et acronymes | xxi |
| Dédicace | xxiii |
| Remerciements | xxv |
| Introduction | 1 |
| Chapitre 1. Notions en économétrie financière | 5 |
| 1.1. Présentation du cadre de la modélisation financière | 5 |
| 1.1.1. Les faits stylisés (<i>stylized facts</i>) des rendements financiers | 6 |
| 1.1.2. Les faits stylisés (<i>stylized facts</i>) des variances réalisées | 6 |
| 1.2. Quelques modèles financiers de la littérature | 7 |
| 1.2.1. Modèle GARCH et ses extensions | 7 |
| 1.2.2. Modèles à volatilité stochastique | 8 |
| 1.2.3. Modèles MEM | 8 |
| 1.2.4. Modèles HAR | 9 |
| 1.3. Les modèles à chaîne de Markov cachée (HMM) | 9 |
| 1.3.1. Estimation des modèles HMM | 11 |
| 1.3.2. Modèles factoriels à chaîne de Markov cachée (FHMM) | 13 |
| 1.3.2.1. Le modèle MSM | 13 |
| 1.3.2.2. Modèle CDRS | 14 |
| 1.3.2.3. Le modèle FHMV | 15 |

| | | |
|--------------------|--|-----------|
| 1.3.3. | Un autre HMM à haute dimension : Le modèle DSARV | 17 |
| 1.4. | L'asymétrie dans les rendements financiers : l'effet levier (<i>leverage effect</i>) | 17 |
| 1.4.1. | Effet levier imbriqué dans la matrice de transition | 18 |
| 1.4.2. | Effet levier multiplicatif | 18 |
| 1.5. | Discussion | 19 |
| Chapitre 2. | Le modèle MDSV : <i>Multifractal Discrete Stochastic Volatility</i> | 21 |
| 2.1. | Cadre général de la modélisation factorielle | 21 |
| 2.1.1. | Présentation du cadre général | 22 |
| 2.1.2. | Quelques propriétés théoriques des modèles factoriels | 24 |
| 2.2. | Le processus MDSV | 26 |
| 2.2.1. | Présentation du processus | 26 |
| 2.2.2. | Interprétations et lien avec la littérature | 28 |
| 2.2.3. | Analyses théoriques et propriétés | 29 |
| 2.3. | Modèles à chaîne de Markov cachée basés sur le processus MDSV | 31 |
| 2.3.1. | Liaison du processus MDSV aux données | 31 |
| 2.3.2. | Estimation et prévision du modèle MDSV | 34 |
| 2.4. | Étude de simulations | 36 |
| 2.5. | L'effet levier (<i>leverage effect</i>) | 39 |
| 2.6. | Discussion | 40 |
| Chapitre 3. | Étude empirique du processus MDSV | 43 |
| 3.1. | Présentation et analyse descriptive des données financières utilisées | 43 |
| 3.2. | Outils de mesures de performance | 45 |
| 3.2.1. | Mesures de performance : ajustement et prévision | 45 |
| 3.2.1.1. | Mesures de performance : ajustement | 46 |
| 3.2.1.2. | Mesures de performance : prévision | 46 |
| 3.2.2. | Mesures de performance : gestion du risque | 48 |
| 3.2.2.1. | Couverture inconditionnelle : Test de Kupiec | 49 |
| 3.2.2.2. | Indépendance : Test de Christoffersen | 49 |
| 3.2.2.3. | Couverture conditionnelle | 50 |

| | |
|---|-----------|
| 3.3. Sélection des nombres de composantes et d'états | 50 |
| 3.3.1. Sélection du nombre d'états | 50 |
| 3.3.2. Sélection du nombre de composantes | 52 |
| 3.4. Modélisation du processus de rendements financiers | 54 |
| 3.4.1. Estimation du modèle | 55 |
| 3.4.2. Performance de prévision du modèle | 55 |
| 3.4.3. Performance de gestion du risque du modèle | 59 |
| 3.5. Modélisation du processus de variances réalisées | 59 |
| 3.5.1. Estimation du modèle | 61 |
| 3.5.2. Performance de prévision du modèle | 62 |
| 3.6. Modélisation conjointe : rendements financiers et variances réalisées | 65 |
| 3.6.1. Estimation du modèle | 66 |
| 3.6.2. Performance de prévision du modèle | 67 |
| 3.6.3. Performance de gestion du risque du modèle | 71 |
| 3.7. MDSV : forme jointe ou forme univariée ? | 72 |
| 3.7.1. Ajustement et prévision des rendements financiers | 72 |
| 3.7.2. Prévision des variances réalisées | 74 |
| 3.7.3. Gestion du risque | 76 |
| 3.8. Analyse de l'ajustement du modèle MDSV joint sur l'indice S&P 500 | 77 |
| 3.8.1. Estimation des paramètres | 78 |
| 3.8.2. Fonction d'autocorrélation | 80 |
| 3.8.3. Analyse de l'effet levier | 81 |
| 3.8.4. Distribution stationnaire de la volatilité | 81 |
| 3.9. Discussion | 83 |
| Chapitre 4. Le modèle iFHMV : <i>infinite Factorial Hidden Markov Volatility</i> | 85 |
| 4.1. Motivation et justification | 85 |
| 4.2. Présentation du processus iFHMV | 86 |
| 4.2.1. Cas constant : Nombre de composantes N fixe | 87 |
| 4.2.2. Nombre infini de composantes : N tendant vers l'infini | 89 |
| 4.2.3. Construction rupture de bâton (<i>stick-breaking</i>) | 91 |

| | |
|---|--------|
| 4.3. Estimation du modèle | 92 |
| 4.3.1. <i>Slice sampling</i> : Estimation de N_+ | 92 |
| 4.3.2. Algorithme d'estimation du modèle iFHMV | 93 |
| 4.4. Analyse empirique du modèle | 96 |
| 4.4.1. Estimation du modèle (<i>in-sample</i>) | 98 |
| 4.4.2. Prévision du modèle (<i>out-of-sample</i>) | 101 |
| 4.4.2.1. La log-vraisemblance prédictive | 101 |
| 4.4.2.2. Les autres mesures de performance de prévision | 102 |
| 4.4.2.3. Nombre de composantes inféré (<i>out-of-sample</i>) | 104 |
| 4.5. Discussion | 104 |
| Conclusion générale | 107 |
| Bibliographie | 111 |
| Annexe A. Annexe relative au chapitre 2 | A-i |
| A.1. Preuve du théorème 2.1.1 | A-i |
| A.2. Preuves de la proposition 2.1.1 | A-iii |
| A.3. Preuve de la proposition 2.1.2 | A-v |
| A.4. Preuve de la proposition 2.2.1 | A-x |
| A.5. Preuve de la proposition 2.2.2 | A-xvi |
| A.6. Preuve de la proposition 2.3.1 | A-xx |
| A.7. Preuve des résultats de prévision | A-xxiv |
| Annexe B. Annexe relative au chapitre 3 | B-i |
| B.1. Estimation et prévision dans le modèle de Liu et Maheu (2018) | B-i |
| B.2. Comparaison de prévision de variances réalisées : forme univariée ou forme jointe | B-i |
| B.3. Estimation du modèle MDSV par maximum de vraisemblance sur plusieurs séries de données | B-iii |
| B.4. Fonctions d'autocorrélation du modèle MDSV sur plusieurs séries de données | B-iii |
| Annexe C. Annexe relative au chapitre 4 | C-i |

| | |
|--|--------|
| C.1. Preuves de la proposition 4.2.1 | C-i |
| C.2. Preuves du théorème 4.2.1..... | C-ii |
| C.3. Preuves de la proposition 4.3.1 | C-v |
| C.4. Les autres distributions <i>a posteriori</i> de la section 4.3.2..... | C-vii |
| C.5. Éléments complémentaires sur les résultats d'estimations | C-viii |

Liste des tableaux

| | | |
|------|--|------|
| 2.1 | Résultats d'estimation du maximum de vraisemblance des simulations Monte Carlo | 37 |
| 2.2 | Sélection des nombres de composantes et d'états par maximum de vraisemblance..... | 39 |
| 3.1 | Statistiques descriptives | 44 |
| 3.2 | Comparaison des performances d'ajustement : rendements financiers | 56 |
| 3.3 | Log-vraisemblance prédictive : rendements financiers..... | 56 |
| 3.4 | Performance de prévision de volatilité : rendements financiers..... | 58 |
| 3.5 | Performance de gestion du risque : rendements financiers..... | 60 |
| 3.6 | Comparaison des performances d'ajustement : variances réalisées..... | 62 |
| 3.7 | Performance de prévision de volatilité : variances réalisées | 64 |
| 3.8 | Comparaison des performances d'ajustement : modèle joint | 67 |
| 3.9 | Log-vraisemblance prédictive : modèle joint | 68 |
| 3.10 | Performance de prévision de volatilité : modèle joint | 70 |
| 3.11 | Performance de gestion du risque : rendements financiers et variances réalisées | 71 |
| 3.12 | Log-vraisemblance et log-vraisemblance prédictive : forme univariée ou forme jointe | 73 |
| 3.13 | Prévision du carré des rendements financiers : forme univariée ou forme jointe..... | 74 |
| 3.14 | Prévision des variances réalisées : forme univariée ou forme jointe (sans effet levier)..... | 75 |
| 3.15 | Gestion du risque : forme univariée ou forme jointe | 76 |
| 3.16 | S&P 500 : Estimation par maximum de vraisemblance | 79 |
| 4.1 | Statistiques descriptives | 97 |
| 4.2 | Comparaison d'estimation | 100 |
| 4.3 | Log-vraisemblance prédictive..... | 102 |
| 4.4 | Performance de prévision de volatilité | 104 |
| B.1 | Comparaison de performance : Modèle de Liu et Maheu (2018) et MDSV..... | B-ii |

| | | |
|-----|--|-------|
| B.2 | Log-vraisemblance prédictive : Modèle de Liu et Maheu (2018) et MDSV | B-ii |
| B.3 | Performance de prévision de volatilité : Modèle de Liu et Maheu (2018) et MDSV | B-iii |
| B.4 | Prévision de variances réalisées : forme univariée ou forme jointe (avec effet levier) | B-iv |
| B.5 | Paramètres estimés par maximum de vraisemblance | B-v |

Table des figures

| | | |
|------|--|----|
| 3.1 | Séries de variances réalisées des indices S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225 sur la période du 1 ^{er} janvier 2000 au 31 décembre 2019. | 45 |
| 3.2 | Évolution de la log-vraisemblance (en ordonnée) selon le nombre d'états K (en abscisse) dans le modèle MDSV sous sa forme jointe lorsque $N = 1$ sur les données financières S&P 500. | 51 |
| 3.3 | Évolution de la log-vraisemblance prédictive, de la quasi-vraisemblance à un pas de temps, de l'écart quadratique moyen de prévision à un pas de temps et de la valeur-p du test de couverture conditionnelle sur la valeur à risque de niveau 95% (en ordonnées) selon le nombre d'états K (en abscisses) dans le modèle MDSV sous sa forme jointe lorsque $N = 1$ sur les données financières de l'indice S&P 500. | 52 |
| 3.4 | Évolution de la log-vraisemblance (en ordonnée) selon le nombre de composantes N (en abscisse) dans le modèle MDSV sous sa forme jointe lorsque $K = 2$ (trait plein), $K = 3$ (tirets), $K = 4$ (pointillés), $K = 5$ (tirets pointillés), $K = 6$ (long tirets), sur les données financières de l'indice S&P 500. | 53 |
| 3.5 | Évolution de la log-vraisemblance prédictive, de la quasi-vraisemblance à un pas de temps, de l'écart quadratique moyen de prévision à un pas de temps et de la valeur-p du test de couverture conditionnelle sur la valeur à risque de niveau 95% (en ordonnées) selon le nombre de composantes N (en abscisses) dans le modèle MDSV sous sa forme jointe lorsque chaque composante a $K = 2$ états sur les données financières de l'indice S&P 500. | 54 |
| 3.6 | Différence entre le BIC des trois spécifications du modèle MDSV et le modèle Real EGARCH en tenant compte de l'effet levier | 68 |
| 3.7 | Valeurs à risque conditionnelles produites sur l'indice S&P 500 par les formes jointe et univariée des trois spécifications du modèle MDSV sur la période du 22 décembre 2016 au 31 décembre 2019 (échantillon de taille 5016). | 77 |
| 3.8 | Fonction d'autocorrélation empirique en noir contre les fonctions d'autocorrélation théoriques des spécifications ($N = 10, K = 2$) en rouge, ($N = 6, K = 3$) en bleu et ($N = 3, K = 10$) en vert du modèle MDSV. | 80 |
| 3.9 | Paramètre l_i | 81 |
| 3.10 | L_t | 81 |

| | | |
|------|--|------|
| 3.11 | Distribution stationnaire de la volatilité (à gauche) et courbe d'impacts de nouvelles (à droite). Les tracés en trait plein, en pointillés et en tirets désignent respectivement les spécifications ($N = 10, K = 2$), ($N = 6, K = 3$) et ($N = 3, K = 10$) du modèle MDSV. À droite, le tracé en tirets et pointillés désigne le modèle Real EGARCH. | 82 |
| 4.1 | Forme <i>lof</i> d'une matrice binaire ayant 20 lignes (désignant les dates) et 37 colonnes (désignant les chaînes de Markov). Les cases grises indiquent que la chaîne de Markov est active. | 89 |
| 4.2 | Construction rupture de bâton. La ligne horizontale noire est de taille 1. A chaque itération, la ligne verticale noire représente le point de rupture, la ligne bleue représente la longueur du bâton, la ligne rouge en pointillée représente le morceau du bâton qui a été jeté. L'expression sous la ligne rouge représente la proportion du bâton qui a été jetée. | 91 |
| 4.3 | Fonctions d'autocorrélation empiriques $\mathbb{C}orr[r_{t+h} , \max(-r_t, 0)]$ en noir, $\mathbb{C}orr[r_{t+h} , \max(r_t, 0)]$ en rouge. La première ligne présente les ACF pour les indices NASDAQ 100 et FTSE 100 et la deuxième ligne présente les ACF pour les actions GES et CS. | 97 |
| 4.4 | Nombre de composantes N_+ inféré par la méthode de l'échantillonnage de Gibbs par blocs en utilisant l'algorithme 2. La première ligne présente les diagrammes à bandes pour les indices NASDAQ 100 et FTSE 100 et la deuxième ligne présente les diagrammes à bandes pour les actions GES et CS. Des informations complémentaires sur les résultats d'estimation sont fournies à l'annexe C.5. | 98 |
| 4.5 | Densités des log-vraisemblances calculées à l'aide des 3 750 dernières itérations de l'estimation des modèles iFHMV (rouge) et iHMM (noir). La première ligne présente les densités des log-vraisemblances pour les indices NASDAQ 100 et FTSE 100 et la deuxième ligne présente les densités des log-vraisemblances pour les actions GES et CS. | 100 |
| 4.6 | Nombre moyen de composantes N_+ inféré pour la prévision sur les 756 jours boursiers. La première ligne présente les graphiques pour les indices financiers NASDAQ 100 et FTSE 100 alors que la deuxième ligne présente les graphiques pour les actions GES et CS. La série de rendements financiers de chaque série est représentée au-dessus du chaque nombre de composantes. | 105 |
| B.1 | Fonction d'autocorrélation empirique en noir contre les fonctions d'autocorrélation théoriques des spécifications ($N = 10, K = 2$) en rouge, ($N = 6, K = 3$) en blue et ($N = 3, K = 10$) en vert du modèle MDSV sur les données des indices S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225. | B-vi |
| C.1 | Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'indices financiers NASDAQ 100. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance. | C-ix |

| | | |
|-----|--|------|
| C.2 | Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'indices financiers FTSE 100. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance. | C-ix |
| C.3 | Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'action GES. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance. | C-x |
| C.4 | Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'actions CS. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance. | C-x |

Liste des sigles et acronymes

| | |
|-----------------|--|
| AFRIMA | AutoRegressive Fractional Integral Moving Average |
| AIC | Akaike Information Criteria |
| AMEX | American Stock Exchange |
| APARCH | Asymmetric Power Generalized AutoRegressive Conditional Heteroskedasticity |
| AR | AutoRegressive |
| ARCH | AutoRegressive Conditional Heteroskedasticity |
| ARMA | AutoRegressive Moving Average |
| BIC | Bayesian Information Criteria |
| CDRS | Component-Driven Regime Switching |
| COVID-19 | Coronavirus Disease 2019 |
| CS | Citigroup Stock |
| DSARV | Discrete Stochastic AutoRegressive Volatility |
| EGARCH | Exponential Generalized AutoRegressive Conditional Heteroskedasticity |
| EM | Expectation-Maximisation |
| FFBS | Forward Filtering Backward Sampling |
| FHMM | Factorial Hidden Markov Models |
| FHMV | Factorial Hidden Markov Volatility |
| FMI | Fonds Monétaire International |
| FSSE | Finite Sample Standard Error |
| FTSE 100 | Financial Times Stock Exchange 100 |
| GARCH | Generalized AutoRegressive Conditional Heteroskedasticity |
| GES | General Electric Stock |
| GJR | Glosten-Jagannathan-Runkle |
| HAR | Heterogeneous AutoRegressive |

| | |
|--------------------|--|
| HAR-RV | Heterogeneous AutoRegressive Realized Volatility |
| HMM | Hidden Markov Models |
| IBP | Indian Buffet Process |
| iFHMM | Infinite Factorial Hidden Markov Models |
| iFHMV | Infinite Factorial Hidden Markov Volatility |
| iHMM | Infinite Hidden Markov Models |
| KSE | Karachi SE 100 |
| MAFE | Mean Absolute Forecast Error |
| MDSV | Multifractal Discrete Stochastic Volatility |
| MEM | Multiplicative Error Models |
| MSM | Markov Switching Multifractal |
| MS-RV | Markov Switching Realized Volatility |
| NASDAQ 100 | National Association of Securities Dealers Automated Quotation 100 |
| NIF | New Impact Function |
| NIKKEI 225 | Nihon Keizai Shinbun 225 |
| NYSE | New York Stock Exchange |
| OMS | Organisation Mondiale de la Santé |
| QLIK | Quasi-Likelihood |
| Real EGARCH | Realized Exponential Generalized AutoRegressive Conditional Heteroskedasticity |
| RMSE | Root Mean Squared Error |
| RMSFE | Root Mean Squared Forecast Error |
| RSD | Relative Standard Deviation |
| RV | Realized Variance |
| S&P 500 | Standard & Poor's 500 |
| SD | Standard Deviation |
| SE | Standard Error |
| SVOL | Stochastic Volatility |
| VAR | Vector AutoRegressive |
| VaR | Value at Risque |

Dédicace

Je dédie cette thèse à mon feu père **Kassimou Maoude**.
Il aurait tant aimé voir l'aboutissement de ce travail.

Remerciements

Cette thèse est l'accomplissement non seulement d'un énorme travail mais également d'un grand soutien de la part des membres de ma famille, de mes professeurs, de mes amis ainsi que celle de plusieurs organismes et fondations encourageant les étudiants. C'est l'occasion pour moi d'exprimer ma gratitude vis-à-vis de toutes les parties qui m'ont apporté ce soutien.

Sur le plan familial et du haut de cette liste, je remercie ma charmante et douce épouse : « Merci chérie pour tout l'amour et le soutien que tu m'apportes chaque jour ». Ensuite, je rends un vibrant hommage à mes parents qui m'ont assisté à chaque étape de ce périple : « C'est grâce à votre éducation et à votre soutien inconditionnel que j'ai pu m'épanouir et continuer à faire des efforts chaque jour ». Merci également à tonton Alassane et tonton Razack pour chacun de vos conseils dans mes moments de doute. Je remercie aussi mes amis et frères Lucien Chaffa, Ismael Assani, Firmin Ayivodji, Aristide Houndetoungan, Marius Adom et Karl-Augustt Alahassa qui m'ont apporté l'environnement familial dont j'avais besoin pour mener à bien mes études : « Votre présence et vos encouragements m'ont été d'une grande aide ».

Sur le plan académique et professionnel, j'ai beaucoup de reconnaissance et d'admiration à témoigner à mes directeurs de recherche Maciej Augustyniak et Mylène Bédard pour leur soutien continu, leurs conseils, leur patience et leurs encouragements. J'aimerais remercier tout particulièrement Maciej Augustyniak qui n'a ménagé aucun effort pour m'assister constamment depuis le début de mes études doctorales. J'ai appris énormément en travaillant avec lui, non seulement du point de vue intellectuel mais également comment vivre, penser et travailler en tant que chercheur. J'aime à penser qu'au fil des années nous sommes devenus de bons amis. C'est l'occasion aussi de remercier tous les professeurs du département de Mathématiques et de Statistique pour les précieux cours qu'ils m'ont donnés et tout le personnel administratif : « Grâce à vous tous, j'ai passé de très bons moments au DMS. Je n'aurais jamais pu imaginer un meilleur environnement d'étude ».

Finalement, je tiens à remercier plusieurs organismes qui m'ont offert un soutien financier pendant mon doctorat : Institut des sciences mathématiques (Bourse d'excellence), Fonds de recherche du québec - Nature et technologies (Bourse de doctorat en recherche), Mitacs (Financement pour stage en entreprise), Université de Montréal (Bourse de scolarité et Bourse Serge Tardif), Centre de recherches mathématiques (financement pour séance de travail en France). Je remercie une fois encore Maciej Augustyniak et Mylène Bédard pour leurs apports financiers.

Introduction

L'année 2020 a été caractérisée par un ensemble de crises dans plusieurs domaines de la vie des populations. Dès janvier, l'Organisation Mondiale de la Santé (OMS) a déclaré la COVID-19 comme étant une urgence de santé publique à portée internationale. Une crise économique et financière a immédiatement accompagné cette crise sanitaire. Le Fond Monétaire International (FMI), qui prévoyait une croissance du revenu par habitant dans 160 de ses 189 pays membres pour 2020, projette une contraction du revenu par habitant dans plus de 170 pays¹ à la suite de ces événements. Par exemple, le taux de croissance du Canada, initialement projeté pour être une hausse (entre 1.6 % et 1.8 % en 2020²) s'est avéré être une baisse (d'un taux annualisé de -2.1% pour le premier trimestre³). Le pétrole a aussi connu une crise sans précédent allant jusqu'à atteindre un prix négatif en avril 2020. Il est évident que la crise du coronavirus a affecté non seulement les habitudes de la population, mais aussi toute l'économie mondiale et en particulier celle du Canada.

Une question d'intérêt pour les statisticiens, économistes, chercheurs, etc. se dégage de cette situation : comment l'étude des données dans le temps peut-elle permettre de détecter différents comportements pendant les périodes de crises ? L'approche naturelle en économétrie est d'étudier et d'analyser les données du passé afin de détecter et de comprendre des changements structurels qui pourraient exister et ainsi faire des prévisions. Aussi, si ces changements structurels sont importants, on voudrait que les modèles aient la flexibilité de proposer des paramètres appropriés selon ces différents états de la nature. Par exemple, l'étude des données financières révèle des caractéristiques particulières empiriquement observées sur les rendements financiers, appelées faits stylisés (*stylized facts*) (Cohen *et al.*, 1972, McCurdy et Morgan, 1987, Rydén *et al.*, 1998, Cont, 2001, Bulla et Bulla, 2006). Une de ces caractéristiques est la lente décroissance de la fonction d'autocorrélation empirique du carré des rendements financiers. Les modèles économétriques basés sur les rendements financiers doivent ainsi reproduire ces caractéristiques le plus fidèlement possible.

Plusieurs modèles financiers ont été proposés dans la littérature pour tenter de reproduire les faits stylisés des données financières. Les modèles de type GARCH (Engle, 1982, Bollerslev, 1986, Nelson, 1991, Glosten *et al.*, 1993, Ding *et al.*, 1993, etc.) sont devenus rapidement populaires puisqu'ils permettent de reproduire

1. Source : Déclaration de la directrice Kristalina Georgieva du FMI le jeudi 9 avril 2020.

2. Source : Rapport Deloitte octobre 2019 <https://www2.deloitte.com/>.

3. Source : Le quotidien de statistique Canada diffusé le 02 mars 2021.

plusieurs faits stylisés des rendements financiers. Cependant, la fonction d'autocorrélation du carré de ces processus décroît exponentiellement. Des modèles à longue mémoire (ex : AFRIMA introduit par [Granger et Joyeux 1980](#) et [Hosking 1981](#), FIGARCH de [Baillie et al. 1996](#)) ont également été utilisés pour modéliser les rendements financiers. Une autre classe de modèles qui a suscité l'intérêt des chercheurs ces dernières décennies est celle des modèles à chaîne de Markov cachée. Suite aux articles de [Hamilton \(1989, 1990\)](#), les modèles à chaîne de Markov cachée sont devenus populaires dans la modélisation des données économiques, actuarielles et financières. Par exemple, [Hardy \(2003\)](#) a recommandé leur utilisation pour gérer et évaluer les risques d'investissement à long terme. Ces modèles ont la particularité de permettre une variation des paramètres selon les états d'une chaîne de Markov latente, permettant ainsi de mieux refléter les changements structurels du processus générateur des données. De plus, les modèles à chaîne de Markov cachée reproduisent raisonnablement bien les faits stylisés observés sur les rendements financiers ([Rydén et al., 1998](#)). Cependant, ces modèles ont de la difficulté à reproduire le niveau élevé de persistance qui est observée sur les données financières et génèrent une fonction d'autocorrélation de décroissance exponentielle. En effet, le temps de séjour d'une chaîne de Markov dans chacun de ses états est distribué selon une loi géométrique.

Historiquement, les modèles à chaîne de Markov ont été estimés avec un faible nombre de régimes (2 à 4 en général, voir [Ang et Bekaert, 2002](#), [Bauwens et al., 2014](#)), car le nombre de paramètres à estimer explose rapidement avec le nombre de régimes et donc l'optimisation devient vite difficile. À titre illustratif, l'estimation d'un modèle à chaîne de Markov cachée classique comptant M états requiert généralement l'estimation d'un minimum de M^2 paramètres. Bien que ce soit une limitation importante de ce type de modèle, l'utilisation d'un faible nombre de régimes permet d'interpréter les états comme les régimes économiques et financiers. Toutefois, [Calvet et Fisher \(2004\)](#) ont introduit une manière élégante de construire un modèle à chaîne de Markov cachée ayant un grand nombre d'états tout en conservant la parcimonie. Leur approche peut se reformuler comme un modèle factoriel à chaîne de Markov cachée (FHMM, *Factorial Hidden Markov Model*). Les modèles factoriels à chaîne de Markov cachée sont caractérisés par une chaîne de Markov ayant un espace d'états de haute dimension (en pratique autour de 1 000 états sont utilisés) construite de manière parcimonieuse en multipliant entre elles des chaînes de Markov indépendantes à espaces d'états de faible dimension. Par exemple, les modèles MSM (*Markov Switching Multifractal* de [Calvet et Fisher, 2004](#)), CDRS (*Component-Driven Regime-Switching* de [Fleming et Kirby, 2013](#)) et FHMV (*Factorial Hidden Markov Volatility* de [Augustyniak et al., 2019](#)) utilisent un processus dépendant du produit de plusieurs chaînes de Markov, appelées composantes, ayant chacune deux états. Ces modèles ont l'avantage de permettre un niveau de persistance élevée et constituent des compétiteurs sérieux pour la classe des modèles de type GARCH.

Dans cette thèse, nous formalisons un cadre général de modélisation factorielle des données financières. Plus spécifiquement, nous proposons un cadre qui englobe les modèles factoriels à chaîne de Markov cachée de la littérature puis nous l'étudions. Ensuite, nous proposons le modèle MDSV (*Multifractal Discrete Stochastic Volatility*) qui généralise les modèles factoriels à chaîne de Markov cachée de la littérature en autorisant les

composantes à avoir plus de deux états. La paramétrisation particulière associée à ce modèle lui accorde suffisamment de flexibilité pour refléter différentes allures des fonctions d'autocorrélation et en particulier celles qui décroissent lentement. En effet, cette paramétrisation fait que le temps de séjour de la chaîne de Markov dans certains de ses états n'est pas distribué selon une loi géométrique. La modélisation par l'entremise des modèles factoriels à chaîne de Markov cachée nécessite qu'on définisse le nombre N de composantes à multiplier et cela suppose qu'il n'existe aucune incertitude sur ce nombre. Cette thèse propose également un algorithme, basé sur une approche bayésienne, qui autorise les données à déterminer le nombre de chaînes de Markov nécessaires à leur modélisation.

Cette thèse est divisée en quatre chapitres. Le chapitre 1 introduit des notions importantes en économétrie financière qui sont nécessaires à la compréhension des chapitres suivants. Dans ce chapitre, nous présentons plus en détail les modèles financiers existants dans la littérature. Le chapitre 2 décrit le cadre de la modélisation factorielle et étudie ses propriétés théoriques. Le modèle MDSV est présenté dans ce chapitre puis est relié aux données. Une analyse empirique détaillée de ce modèle sur les données de 31 indices financiers est effectuée dans le chapitre 3. Dans le chapitre 4, nous définissons un modèle factoriel à chaînes de Markov cachée obtenu en multipliant un nombre infini de chaînes de Markov à deux états : le modèle iFHMV (*infinite Factorial Hidden Markov Volatility*). En s'inspirant du processus du buffet indien (IBP, *Indian Buffet Process*), nous proposons un algorithme afin d'estimer ce modèle. Nous concluons cette thèse en proposant certaines avenues de recherche future.

Chapitre 1

Notions en économétrie financière

Ce chapitre présente une introduction à la modélisation des données financières et définit également les notions et outils d'économétrie financière nécessaires à la compréhension des chapitres suivants. À l'aide de ces concepts, il justifie finalement l'intérêt des modèles à chaîne de Markov cachée auxquels nous ferons allusion tout au long de cette thèse.

1.1. Présentation du cadre de la modélisation financière

La modélisation financière s'effectue à l'aide de processus financiers (processus de prix, rendements, variances ou volatilités réalisées, etc.) liés entre eux et permettant de mesurer l'activité financière. Dans cette section, nous ferons une présentation générale de la modélisation des données financières.

Soit $\{y_t\}_{t \in \mathbb{R}^+}$, un processus stochastique représentant l'évolution du prix d'un actif financier (action, indice financier ou taux de change), où l'indice t est exprimé en jours. Nous nous intéressons premièrement à la modélisation du log-rendement (*log-return*) sur cet actif. Le log-rendement sur la période $[t-1, t]$ est noté r_t et est défini par :

$$r_t = \log y_t - \log y_{t-1}.$$

Outre le log-rendement, nous étudierons aussi la variance réalisée (*realized variance*) de cet actif financier, notée RV_t et définie par :

$$RV_t = \sum_{i=1}^m r_{t,i}^2, \quad \text{avec } r_{t,i} = \log y_{t+\frac{i}{m}} - \log y_{t+\frac{i-1}{m}},$$

où $r_{t,i}$ est un log-rendement à haute fréquence dans la journée t couvrant la période $[t + \frac{i-1}{m}, t + \frac{i}{m}]$ (par exemple, des log-rendements aux cinq minutes entre l'ouverture et la fermeture de la bourse). Nous notons aussi par rv_t la volatilité réalisée (*realized volatility*), définie par :

$$rv_t = \sqrt{RV_t}.$$

Dans la suite de ce document, nous considérons des processus à temps discret sur l'ensemble d'indices $\{1, 2, \dots, T\}$ où T représente la taille de l'échantillon. Nous désignerons aussi par rendements financiers la

série des log-rendements centrés c'est-à-dire la série

$$r_t - \frac{1}{T} \sum_{t=1}^T r_t.$$

1.1.1. Les faits stylisés (*stylized facts*) des rendements financiers

Les données de rendements financiers présentent des caractéristiques très intéressantes et difficiles à reproduire, mais empiriquement observées (Cohen *et al.*, 1972, McCurdy et Morgan, 1987, Rydén *et al.*, 1998, Cont, 2001, Bulla et Bulla, 2006). Ainsi, un modèle économétrique basé sur les rendements financiers devrait pouvoir reproduire ces caractéristiques, qui sont :

- (a) *Absence d'autocorrélation linéaire dans les rendements* : La fonction d'autocorrélation échantillonnale définie par $\text{Corr}[r_{t+h}, r_t]$ est approximativement nulle pour tout $h = 1, 2, \dots$
- (b) *Présence d'autocorrélation positive dans le processus $\{|r_t|^p\}_{p>0}$* : La fonction d'autocorrélation échantillonnale des processus $\{r_t^2\}$ et $\{|r_t|\}$ est positive à $h = 1$, puis décroît lentement lorsque h augmente. Cela implique une certaine dépendance temporelle dans les données et une persistance de la volatilité.
- (c) *Regroupement des volatilités* : Les rendements de faibles et de fortes volatilités tendent à être regroupés.
- (d) *Queues épaisses et asymétrie négative* : Le coefficient d'asymétrie (*skewness*) de la densité des rendements financiers est généralement négatif et son coefficient d'aplatissement (*kurtosis*) est généralement plus élevé que celui de la loi normale qui est 3. On dit que les rendements financiers sont à queues épaisses avec une asymétrie négative.
- (e) *Effet levier* : Un choc négatif significatif sur les rendements financiers au temps t tend à accroître la volatilité des rendements subséquents par un plus grand écart comparativement à un choc positif de la même amplitude.

1.1.2. Les faits stylisés (*stylized facts*) des variances réalisées

Les données de variances réalisées et de volatilités réalisées présentent elles aussi des caractéristiques particulières (Ebens *et al.*, 1999, Andersen *et al.*, 2001, Huang *et al.*, 2007). Ainsi, un modèle économétrique basé sur les variances réalisées devrait pouvoir reproduire ces caractéristiques, qui sont :

- (a) *Non constance et regroupement des volatilités* : Les volatilités réalisées varient dans le temps et leur valeurs tendent à être regroupées dans le temps.
- (b) *Normalité des rendements standardisés $\frac{r_t}{rv_t}$* : Bien que la distribution des rendements soit non-gaussienne et leptokurtique, la distribution des rendements standardisés (le rapport du rendement à la volatilité réalisée) se rapproche d'une distribution normale.
- (c) *Distribution log-normale des variances réalisées* : La distribution du logarithme des variances réalisées est proche de la normale.

- (d) *Dépendance à long terme* : Le processus de variances réalisées présente une dépendance à long terme. En effet, le paramètre de mémoire du modèle fractionnel intégré (voir modèle AFRIMA de [Baillie, 1996](#)) est généralement plus grand que 0.
- (e) *Stationnarité des variances réalisées* : Les tests de racine unitaire standards rejettent souvent la présence d'une racine unitaire dans les variances et les volatilités réalisées quotidiennes, ce qui atteste de la stationnarité de ces processus.
- (f) *Asymétrie de la fonction d'impact des nouvelles (News Impact Function)* : La fonction d'impact des nouvelles (NIF) définie par [Engle et Ng \(1993\)](#) est asymétrique et en forme de V.

1.2. Quelques modèles financiers de la littérature

Au cours des dernières décennies, plusieurs modèles financiers ont été proposés pour reproduire les faits stylisés spécifiques aux rendements financiers et aux variances réalisées. Dans la plupart de ces modèles, les rendements financiers sont modélisés par :

$$r_t = \mu_t + \sqrt{V_t} \epsilon_t,$$

où $\{\mu_t\}$ est un processus de moyenne conditionnelle, $\{V_t\}$ un processus de variance conditionnelle et $\{\epsilon_t\}$ représente un processus d'innovation. Dans le cadre de cette thèse, nous nous intéressons principalement au processus de variance conditionnelle $\{V_t\}$ car nous étudions la série centrée des log-rendements. Ainsi, l'équation,

$$r_t = \sqrt{V_t} \epsilon_t, \tag{1.2.1}$$

sera généralement utilisée pour la plupart des modèles présentés dans cette thèse.

Nous présentons dans cette section quelques modèles financiers de la littérature. Plusieurs de ces modèles seront utilisés dans les chapitres suivants comme des modèles de référence pour mesurer la performance des modèles que nous proposerons.

1.2.1. Modèle GARCH et ses extensions

[Engle \(1982\)](#) a introduit le modèle ARCH (*AutoRegressive Conditional Heteroscedasticity*) : un processus non corrélé de moyenne nulle, de variance conditionnelle (par rapport au passé) non constante et de variance inconditionnelle constante. La généralisation naturelle du modèle ARCH fut proposée par [Bollerslev \(1986\)](#) pour tenir compte des variances conditionnelles passées dans l'équation de variance conditionnelle actuelle. Ces processus et leurs extensions (GJR GARCH de [Glosten et al., 1993](#), APARCH de [Ding et al., 1993](#), EGARCH de [Nelson, 1991](#), etc.) sont devenus très populaires dans la modélisation des rendements financiers puisqu'ils permettent de reproduire plusieurs faits stylisés.

Le modèle ARCH(q) de [Engle \(1982\)](#) modélise la variance conditionnelle V_t par l'entremise de la relation suivante :

$$V_t = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2, \tag{1.2.2}$$

où $\omega > 0$ et $\alpha_i \geq 0, i = 1, \dots, q$. Le modèle GARCH(1, 1), quant à lui, modélise le processus de variance conditionnelle $\{V_t\}$ comme suit :

$$V_t = \omega + \alpha \epsilon_{t-1}^2 + \beta V_{t-1}, \quad (1.2.3)$$

où $\alpha, \beta \geq 0$. Ce modèle GARCH(1,1) permet une autocorrélation du processus $\{\epsilon_t^2\}$, ce qui génère des périodes de faibles et de fortes volatilités. De plus, lorsque $\alpha + \beta < 1$, le processus de rendements $\{r_t\}$ est stationnaire au sens strict (*strictly stationary*) et au second ordre. Dans ce cas, la variance inconditionnelle est donnée par :

$$\text{Var}[r_t] = \frac{\omega}{1 - \alpha - \beta}.$$

1.2.2. Modèles à volatilité stochastique

Le modèle à volatilité stochastique (SVOL, *Stochastic Volatility*) proposé par Andersen (1994) est l'alternatif naturel aux modèles de la famille ARCH. Il autorise les moyenne et variance conditionnelles à être dirigées par des processus d'innovations différents. Fridman et Harris (1998) montrent que le modèle à volatilité stochastique apporte une flexibilité supplémentaire comparativement aux modèles de la famille ARCH. Ce modèle SVOL, en temps discret, est défini comme suit :

$$\begin{cases} r_t &= \sqrt{V_t} \epsilon_t, \\ \log V_t &= \beta_0 + \beta_1 \log V_{t-1} + \sigma_\nu \nu_t, \end{cases}$$

où $\beta_0, \beta_1, \sigma_\nu$ sont des paramètres du modèle. Dans le modèle de Fridman et Harris (1998), les innovations ϵ_t et ν_t sont indépendantes et suivent une distribution normale de moyenne nulle et de variance 1. Une extension multicomposantes de ce modèle a été proposée par Andersen et Bollerslev (1997). Elle suppose que la volatilité V_t est obtenue à partir de N ($N \in \mathbb{N}$) composantes $V_t^{(i)}, i = 1, 2, \dots, N$:

$$\begin{cases} V_t &= \sigma^2 \exp\left(\sum_{i=1}^N \log V_t^{(i)}\right), \\ \log V_t^{(i)} &= \beta^{(i)} \log V_{t-1}^{(i)} + \sigma_\nu^{(i)} \nu_t^{(i)}, \quad i = 1, 2, \dots, N, \end{cases}$$

où $\sigma^2, \beta^{(i)}, \sigma_\nu^{(i)}$ sont des paramètres du modèle et les innovations $\nu_t^{(i)}, i = 1, 2, \dots, N$ sont indépendantes entre elles et indépendantes de ϵ_t .

1.2.3. Modèles MEM

La classe des modèles ARCH a principalement été introduite pour modéliser les rendements financiers. Étant donné que le processus d'innovation a normalement un support sur \mathbb{R} dans ces modèles, ils ne sont pas directement applicables pour modéliser des processus positifs tels que les variances réalisées, le volume d'actions échangées sur des périodes de 10 minutes ou encore le prix le plus élevé moins le prix le plus bas sur une période de temps spécifique (*high-low range*). Le modèle MEM (*Multiplicative Error Model*) a été introduit par Engle (2002) pour répondre à cette problématique. Si RV_t est la variable à modéliser, alors le modèle se présente comme suit :

$$RV_t = V_t \eta_t, \quad (1.2.4)$$

où $\{\eta_t\}$ est un processus d'innovation positif vérifiant $\mathbb{E}[\eta_t] = 1$. Cette définition peut se combiner avec la classe des modèles ARCH puisqu'on peut appliquer n'importe quelle spécification ARCH pour modéliser le processus $\{V_t\}$. Les modèles MEM sont rapidement devenus populaires et ont été appliqués à plusieurs séries de données financières positives (De Luca *et al.*, 2010, Brownlees *et al.*, 2011, Engle *et al.*, 2012, Cipollini *et al.*, 2013, Augustyniak *et al.*, 2019). Cependant, d'autres modèles ont également été proposés pour modéliser les variances réalisées; le modèle HAR en est un exemple.

1.2.4. Modèles HAR

Le modèle HAR-RV (*Heterogeneous AutoRegressive Realized Volatility*) proposé par Corsi (2009) est conçu comme une spécification additive de la variance réalisée agrégée à différents horizons temporels. Une extension directe, le modèle log-HAR développé par Corsi *et al.* (2009) et Corsi et Renò (2012), modélise plutôt le logarithme des variances réalisées. Ce modèle log-HAR est défini comme suit :

$$\log RV_t = \beta_0 + \beta_1 \log RV_t^{(1)} + \beta_2 \log RV_t^{(5)} + \beta_3 \log RV_t^{(22)} + \epsilon_t, \quad (1.2.5)$$

où $\beta_0, \beta_1, \beta_2, \beta_3$ sont des paramètres du modèle, ϵ_t suit une distribution normale de moyenne nulle et de variance σ_ϵ^2 et les variables explicatives sont définies par $\log RV_t^{(h)} = \frac{1}{h} \sum_{i=1}^h \log RV_{t-i}$ pour $h = 1, 5$ et 22 . Des extensions de ce modèle ont été proposées par Andersen *et al.* (2007) et Bauwens *et al.* (2012) pour prendre en compte l'effet levier et les sauts (des changements abrupts qui pourraient être présents dans les données). Le modèle log-HAR peut être représenté sous une forme MEM. En effet, l'équation (1.2.5) peut être réexprimée comme

$$RV_t = V_t \eta_t,$$

où

$$V_t = e^{\beta_0} \left(RV_t^{(1)}\right)^{\beta_1} \left(RV_t^{(5)}\right)^{\beta_2} \left(RV_t^{(22)}\right)^{\beta_3}$$

et

$$\eta_t = \exp(\epsilon_t - \sigma_\epsilon^2/2).$$

1.3. Les modèles à chaîne de Markov cachée (HMM)

Les modèles à chaîne de Markov cachée permettent de modéliser le processus de variance conditionnelle $\{V_t\}$ de l'équation (1.2.1) autrement que par un processus prévisible comme dans le modèle GARCH. Un modèle à chaîne de Markov cachée peut être approximativement vu comme une chaîne de Markov observée de façon imparfaite et bruitée sous forme de données. En effet, les modèles à chaîne de Markov cachée (HMM, *Hidden Markov Models*), aussi appelés modèles à changement de régimes markoviens (*Markov Switching Models*), sont des modèles faisant intervenir un processus stochastique non-observé, noté $\{C_t\}$ (une chaîne de Markov à temps discret et à espace d'états discret de dimension finie). Ce processus latent régit la distribution d'un processus observé, noté $\{y_t\}$, qui peut être univarié (comme par exemple le rendement d'un actif financier) ou multivarié (comme par exemple le couple de variables « rendement et variance

réalisée » d'un actif financier). Mathématiquement, un modèle à chaîne de Markov cachée est un processus $\{y_t, C_t\}_{t \in \mathbb{N}}$ où la densité de y_t conditionnellement à y_{t-1}, y_{t-2}, \dots et C_t, C_{t-1}, \dots ne dépend que de l'état de la chaîne de Markov $\{C_t\}$ à la date t . Avant de définir concrètement le modèle à chaîne de Markov cachée, voici quelques définitions importantes :

Définition 1.3.1. Une chaîne de Markov $\{C_t\}$ à temps discret et à espace d'états discret et fini, est un processus à temps discret prenant ses valeurs dans un ensemble fini E (appelé espace d'états) et vérifiant la propriété de Markov suivante :

$$\forall t, \quad \mathbb{P}(C_t = c_t \mid C_{t-1} = c_{t-1}, \dots, C_0 = c_0) = \mathbb{P}(C_t = c_t \mid C_{t-1} = c_{t-1}),$$

où $c_0, c_1, \dots, c_t \in E$. Dans le cadre de cette thèse, on désignera par « dimension de la chaîne de Markov $\{C_t\}$ », la cardinalité de son espace d'états E , notée $|E|$. Pour alléger la notation dans ce qui suit, nous supposons sans perte de généralité que $E = \{1, 2, \dots, M\}$.

Définition 1.3.2. Une chaîne de Markov $\{C_t\}$ à temps discret et à espace d'états discret de dimension finie $M \in \mathbb{N}$, est dite homogène (time-homogeneous) lorsque les probabilités de transition d'un état $C_{t-1} = i$ à un état $C_t = j$ sont indépendantes du temps, c'est-à-dire,

$$\forall t, \quad \mathbb{P}(C_t = j \mid C_{t-1} = i) = p_{ij},$$

où $0 \leq p_{ij} \leq 1$ et $i, j \in E$.

Définition 1.3.3. On appelle matrice de transition de la chaîne de Markov $\{C_t\}$, la matrice

$$\mathbf{P}_t = [\mathbb{P}(C_t = j \mid C_{t-1} = i)]_{i, j \in E},$$

regroupant toutes les probabilités de transition d'un état à un autre. Si la chaîne $\{C_t\}$ est homogène alors \mathbf{P}_t est indépendante du temps et

$$\mathbf{P}_t = \mathbf{P} = [p_{ij}]_{i, j \in E}.$$

Il est souvent supposé que la chaîne de Markov est homogène, irréductible et apériodique. Ces hypothèses sont émises d'abord par souci de simplicité, mais aussi pour éviter des situations où la chaîne se retrouverait piégée dans un sous-ensemble de l'espace d'états. Ces suppositions nous permettent d'avoir une chaîne de Markov ergodique, donc vérifiant :

$$\lim_{t \rightarrow \infty} \mathbb{P}(C_t = j \mid C_0 = i) = \pi_j, \quad \forall i, j \in E,$$

où $\{\pi_j\}_{j \in E}$ est un vecteur de probabilités correspondant à la loi stationnaire de la chaîne de Markov $\{C_t\}$.

Définition 1.3.4. On appelle temps de séjour de la chaîne de Markov $\{C_t\}$ dans l'état i , la variable aléatoire $S(i)$ à valeur dans \mathbb{N} représentant la durée totale de séjour dans l'état i avant le premier passage dans un autre état sachant que l'état initial est i .

Remarque 1.3.1. Le temps de séjour de la chaîne de Markov $\{C_t\}$ dans l'état i est distribuée selon une loi géométrique de probabilité de succès p_{ii} . Pour tout $k = 0, 1, 2, \dots$,

$$\mathbb{P}(S(i) = k) = \mathbb{P}(C_{k+1} \neq i, C_k = i, C_{k-1} = i, \dots, C_1 = i \mid C_0 = i) = p_{ii}^k (1 - p_{ii}).$$

Un modèle à chaîne de Markov cachée est donc explicitement défini comme suit :

Définition 1.3.5. *Un modèle à chaîne de Markov cachée est un processus multivarié à temps discret partiellement observé, noté $\{y_t, C_t\}$, vérifiant :*

$$\left\{ \begin{array}{l} C_1 \sim p_\theta(c_1) \\ C_t | (y_{1:t-1}, C_{1:t-1} = c_{1:t-1}) \sim p_\theta(c_t | c_{t-1}), \quad t = 2, 3, \dots \\ y_t | (y_{1:t-1}, C_{1:t} = c_{1:t}) \sim p_\theta(y_t | c_t), \quad t = 1, 2, \dots, \end{array} \right.$$

où $y_{1:t-1}$ et $C_{1:t}$ sont respectivement $\{y_1, y_2, \dots, y_{t-1}\}$ et $\{C_1, C_2, \dots, C_t\}$, θ représente l'ensemble des paramètres du modèle à estimer et $p_\theta(\cdot)$ désigne la loi de probabilité étant donné θ .

Une définition plus générale est parfois utilisée. Elle autorise la matrice de transition de la chaîne de Markov cachée à être non-homogène et que le processus $\{y_t\}$ ne soit plus nécessairement une suite de variables aléatoires indépendantes conditionnellement aux valeurs prises par la chaîne de Markov $\{C_t\}$. Elle stipule qu'un modèle à chaîne de Markov $\{y_t, C_t\}$ vérifie

$$\left\{ \begin{array}{l} C_1 \sim p_\theta(c_1) \\ C_t | (y_{1:t-1}, C_{1:t-1} = c_{1:t-1}) \sim p_\theta(c_t | c_{t-1}, y_{1:t-1}), \quad t = 2, 3, \dots \\ y_t | (y_{1:t-1}, C_{1:t} = c_{1:t}) \sim p_\theta(y_t | c_t, y_{1:t-1}), \quad t = 1, 2, \dots, \end{array} \right.$$

où θ représente l'ensemble des paramètres du modèle à estimer et $p_\theta(\cdot)$ désigne la loi de probabilité étant donné θ .

1.3.1. Estimation des modèles HMM

Soit un modèle à chaîne de Markov cachée $\{y_t, C_t\}$. L'estimation de ce modèle peut être accomplie par le maximum de vraisemblance à travers deux techniques existantes : le filtre d'Hamilton et l'algorithme espérance-maximisation (EM). Dans cette section, nous présenterons le filtre d'Hamilton (*Hamilton filter*) alternativement appelé l'algorithme de filtrage avant (*forward filtering algorithm*).

Le filtre d'Hamilton est un algorithme proposé par [Hamilton \(1989, 1990\)](#) pour calculer la log-vraisemblance des modèles à chaîne de Markov cachée.

Proposition 1.3.1. *Soit $\{y_t, C_t\}_{t \in \mathbb{N}}$ un modèle à chaîne de Markov cachée défini par les équations suivantes :*

$$\left\{ \begin{array}{l} C_1 \sim p_\theta(c_1) \\ C_t | (y_{1:t-1}, C_{1:t-1} = c_{1:t-1}) \sim p_\theta(c_t | c_{t-1}, y_{1:t-1}), \quad t = 2, 3, \dots \\ y_t | (y_{1:t-1}, C_{1:t} = c_{1:t}) \sim p_\theta(y_t | c_t, y_{1:t-1}), \quad t = 1, 2, \dots, \end{array} \right.$$

La log-vraisemblance d'un échantillon (possiblement multivarié) $y_{1:T}$ provenant de ce modèle est donnée par

$$\mathcal{L} = \log p_\theta(y_1) + \sum_{t=2}^T \log p_\theta(y_t | y_{1:t-1}),$$

et peut être calculée à l'aide de l'algorithme de filtrage avant suivant.

Algorithme 1 Algorithme de filtrage avant

1: Calculer la vraisemblance de y_1 et la distribution de C_1 conditionnellement à y_1 :

$$p_\theta(y_1) = \sum_{j=1}^M p_\theta(C_1 = j) p_\theta(y_1 | C_1 = j)$$

et

$$p_\theta(C_1 = j | y_1) = \frac{p_\theta(C_1 = j) p_\theta(y_1 | C_1 = j)}{p_\theta(y_1)}, \quad j \in E;$$

2: Pour $t = 2, 3, \dots, T$, calculer la vraisemblance de y_t conditionnellement à $y_{1:t-1}$ et la distribution de C_t conditionnellement à $y_{1:t}$:

$$p_\theta(C_t = j | y_{1:t-1}) = \sum_{i=1}^M p_\theta(C_t = j | C_{t-1} = i, y_{1:t-1}) p_\theta(C_{t-1} = i | y_{1:t-1});$$

$$p_\theta(y_t | y_{1:t-1}) = \sum_{j=1}^M p_\theta(y_t | y_{1:t-1}, C_t = j) p_\theta(C_t = j | y_{1:t-1});$$

$$p_\theta(C_t = j | y_{1:t}) = \frac{p_\theta(C_t = j | y_{1:t-1}) p_\theta(y_t | y_{1:t-1}, C_t = j)}{p_\theta(y_t | y_{1:t-1})}, \quad j \in E;$$

3: Calculer la vraisemblance du modèle par la formule :

$$\mathcal{L} = \log p_\theta(y_{1:T}) = \log p_\theta(y_1) + \sum_{t=2}^T \log p_\theta(y_t | y_{1:t-1}).$$

DÉMONSTRATION. En utilisant les formules de probabilité totale et de Bayes on obtient successivement :

$$p_\theta(y_1) = \sum_{j=1}^M p_\theta(y_1, C_1 = j) = \sum_{j=1}^M p_\theta(C_1 = j) p_\theta(y_1 | C_1 = j),$$

et

$$p_\theta(C_1 = j | y_1) = \frac{p_\theta(y_1, C_1 = j)}{p_\theta(y_1)} = \frac{p_\theta(C_1 = j) p_\theta(y_1 | C_1 = j)}{p_\theta(y_1)}, \quad j \in E.$$

Par la suite, pour $t = 2, \dots, T$, on calcule récursivement :

$$\begin{aligned} p_\theta(C_t = j | y_{1:t-1}) &= \sum_{i=1}^M p_\theta(C_t = j, C_{t-1} = i | y_{1:t-1}) \\ &= \sum_{i=1}^M p_\theta(C_t = j | C_{t-1} = i, y_{1:t-1}) p_\theta(C_{t-1} = i | y_{1:t-1}) \end{aligned}$$

et

$$\begin{aligned}
p_\theta(y_t | y_{1:t-1}) &= \sum_{j=1}^M p_\theta(y_t, C_t = j | y_{1:t-1}) \\
&= \sum_{j=1}^M p_\theta(y_t | y_{1:t-1}, C_t = j) p_\theta(C_t = j | y_{1:t-1}) \\
&= \sum_{j=1}^M \left(p_\theta(y_t | y_{1:t-1}, C_t = j) \sum_{i=1}^M p_\theta(C_t = j | C_{t-1} = i, y_{1:t-1}) p_\theta(C_{t-1} = i | y_{1:t-1}) \right).
\end{aligned} \tag{1.3.1}$$

Ainsi

$$p_\theta(C_t = j | y_{1:t}) = \frac{p_\theta(y_t, C_t = j | y_{1:t-1})}{p_\theta(y_t | y_{1:t-1})} = \frac{p_\theta(C_t = j | y_{1:t-1}) p_\theta(y_t | y_{1:t-1}, C_t = j)}{p_\theta(y_t | y_{1:t-1})}, \quad j \in E.$$

Finalement, la log-vraisemblance s'obtient avec la relation suivante :

$$\mathcal{L} = \log p_\theta(y_{1:T}) = \log p_\theta(y_1) + \sum_{t=2}^T \log p_\theta(y_t | y_{1:t-1}).$$

□

Cette fonction de log-vraisemblance peut être maximisée numériquement en fonction des paramètres du modèle.

1.3.2. Modèles factoriels à chaîne de Markov cachée (FHMM)

Un modèle factoriel à chaîne de Markov cachée (FHMM) est un modèle dans lequel on multiplie plusieurs chaînes de Markov, généralement à espaces d'états de faibles dimensions, pour construire une chaîne à espace d'états de grande dimension (Ghahramani et Jordan, 1997). Cette famille de modèles offre la possibilité d'obtenir une chaîne de Markov à espace d'états de grande dimension en utilisant très peu de paramètres. Nous présentons, dans la suite de cette section, les modèles MSM (*Markov Switching Multifractal*) de Calvet et Fisher (2004), CDRS (*Component-Driven Regime-Switching*) de Fleming et Kirby (2013) et FHMV (*Factorial Hidden Markov Volatility*) de Augustyniak et al. (2019), qui sont tous des modèles factoriels à chaîne de Markov cachée appliqués au domaine financier. Ces auteurs proposent de modéliser la série des rendements financiers par l'équation (1.2.1) où le processus de variance conditionnelle $\{V_t\}$ est une chaîne de Markov de haute dimension construite à l'aide du produit de plusieurs chaînes de Markov de dimension 2.

1.3.2.1. Le modèle MSM

Dans le modèle MSM proposé par Calvet et Fisher (2004), les rendements financiers satisfont :

$$r_t = \sqrt{V_t} \epsilon_t,$$

où ϵ_t est de distribution gaussienne standard et V_t représente une volatilité stochastique obtenue par le produit de N composantes de volatilité $C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)}$. De manière spécifique,

$$V_t = \sigma^2 \prod_{i=1}^N C_t^{(i)},$$

où σ^2 est la variance inconditionnelle du rendement r_t sous l'hypothèse que les composantes $C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)}$ sont indépendantes. Chacune des composantes est donc définie de sorte à être aléatoire, positive, persistante et d'espérance unitaire de la manière suivante :

$$\begin{cases} C_t^{(i)} \text{ tiré d'une loi } \mathcal{C} & \text{avec probabilité } \gamma_i \\ C_t^{(i)} = C_{t-1}^{(i)} & \text{avec probabilité } 1 - \gamma_i \end{cases},$$

où $\gamma_i \in (0, 1)$, $i = 1, 2, \dots, N$. Le modèle MSM de [Calvet et Fisher \(2004\)](#) autorise une grande variété de spécifications selon le choix de la loi \mathcal{C} ; c'est cependant la spécification binomiale qui est la plus répandue. Dans cette spécification, les composantes $C_t^{(i)}$ sont des chaînes de Markov à deux états ν_0 et $2 - \nu_0$, avec $\nu_0 \in (0, 1)$, qui correspondent aux deux valeurs prises par la loi binomiale \mathcal{C} avec probabilité $\frac{1}{2}$. Les fréquences des composantes de volatilité permettant d'obtenir les probabilités de transition sont liées par la formule :

$$\forall i, \quad \gamma_i = 1 - (1 - \gamma_1)^{b^{i-1}},$$

où $\gamma_1 \in (0, 1)$ et $b > 1$. Par sa construction, le modèle MSM permet des changements de régime à basse fréquence, et donc de longs cycles de volatilité dans les trajets d'échantillons. Les processus estimés génèrent aussi des cycles de volatilité avec des périodes proportionnelles à la taille de l'échantillon, une propriété également apparente dans les chemins d'échantillonnage des processus à mémoire longue. En effet, à travers leur modèle, [Calvet et Fisher \(2004\)](#) ont développé un mécanisme unique qui capture trois caractéristiques généralement considérées comme distinctes dans la littérature : les variations de basse fréquence, la dynamique des fréquences intermédiaires généralement attribuée aux transitions douces autorégressives et les commutateurs hautes fréquences générant des valeurs aberrantes substantielles. Un autre avantage du modèle MSM est sa parcimonie.

1.3.2.2. *Modèle CDRS*

Gardant la même structure factorielle que le modèle MSM, le modèle CDRS proposé par [Fleming et Kirby \(2013\)](#) introduit des changements significatifs sur l'espace d'états et la paramétrisation des matrices de transition, entraînant donc des implications différentes sur la dynamique du processus.

Les auteurs supposent, comme [Calvet et Fisher \(2004\)](#), que la chaîne de Markov qui décrit la dynamique des régimes peut être décomposée en un certain nombre de chaînes de Markov à deux états évoluant indépendamment dans le temps. Ils utilisent cependant le fait qu'une chaîne de Markov à deux états a une représentation autorégressive de premier ordre (AR(1)). Avec cette représentation et le fait que la somme de processus AR(1) indépendants est un processus autorégressif moyenne mobile (ARMA), ils obtiennent un modèle à plusieurs composantes permettant de décrire des dynamiques complexes de volatilité. Si $E = \{1, 2\}$

est l'espace d'états d'une chaîne de Markov, $\mathbf{P} = (p_{ij})_{i,j \in E}$ sa matrice de transition et qu'on désigne par $\mathbf{x}_t = \begin{pmatrix} x_{1,t} & x_{2,t} \end{pmatrix}'$ le vecteur 2×1 indiquant l'état (la j -ième composante de \mathbf{x}_t est égale à 1 et l'autre composante est nulle lorsque la chaîne de Markov est dans l'état j à la date t), alors

$$\mathbf{x}_{t+1} = \mathbf{P}'\mathbf{x}_t + \mathbf{e}_{t+1}, \quad t = 1, \dots, T,$$

où $\{\mathbf{e}_t\}$ est un processus tel que $\mathbb{E}[\mathbf{e}_t | \mathbf{x}_{t-1}, \dots, \mathbf{x}_1] = \mathbf{0}$. Cette équation peut être transformée en utilisant le fait que $x_{2,t} = 1 - x_{1,t}$ comme suit

$$x_{1,t+1} = \omega + \phi(x_{1,t} - \omega) + e_{1,t+1}, \quad t = 1, \dots, T,$$

où $\phi = p_{11} - p_{21}$ et $\omega = p_{21}/(1 - p_{11} + p_{21})$ et $e_{1,t+1}$ désigne la première composante de \mathbf{e}_t . Cette représentation est celle d'un processus autorégressif de premier ordre (AR(1)).

Considérons N processus autorégressifs $(x_{1,t}^{(i)})$, $i = 1, \dots, N$, tels que chaque processus $(x_{1,t}^{(i)})$ est régi par l'équation

$$x_{1,t+1}^{(i)} = \omega_i + \phi^{(i)}(x_{1,t}^{(i)} - \omega_i) + e_{1,t+1}^{(i)}, \quad t = 1, \dots, T.$$

Les auteurs supposent que chacune de ces composantes a une contribution additive au logarithme du processus de variance conditionnelle $\{V_t\}$. Le modèle CDRS se présente donc comme suit

$$\log V_t = \beta_0 + \beta_1 \sum_{i=1}^N x_{1,t}^{(i)}, \quad (1.3.2)$$

où $\beta_1 \geq 0$. Les auteurs créent ainsi un effet multiplicatif entre les chaînes de Markov. En effet, de l'équation (1.3.2) on obtient :

$$V_t = \exp\left(\beta_0 + \beta_1 \sum_{i=1}^N x_{1,t}^{(i)}\right) = \sigma^2 \prod_{i=1}^N C_t^{(i)},$$

où $\sigma^2 = \exp(\beta_0)$ et $C_t^{(i)} = \exp(\beta_1 x_{1,t}^{(i)})$. Ce modèle implique ainsi une représentation factorielle de $\{V_t\}$ construite à partir de la multiplication de chaînes de Markov de dimension 2. [Fleming et Kirby \(2013\)](#) créent ainsi une chaîne de Markov de dimension 2^N en utilisant $2N$ paramètres $(\omega_1, \phi^{(1)}, \dots, \omega_N, \phi^{(N)})$ pour définir la matrice de transition. Afin d'obtenir un modèle plus parcimonieux, une spécification qu'ils désignent par CDRS($N, N, N - 1$) consiste à imposer les contraintes $\omega_1 = \omega_2 = \dots = \omega_N$. La matrice de transition de la chaîne de Markov est ainsi construite à partir de $N + 1$ paramètres et les 2^N états sont constitués de $N + 1$ valeurs différentes.

1.3.2.3. Le modèle FHMV

Le modèle FHMV proposé par [Augustyniak et al. \(2019\)](#) conserve une structure factorielle similaire à celle du MSM et du CDRS, mais introduit des changements significatifs sur l'espace d'états en le rendant plus riche. Le modèle FHMV se présente comme le produit d'une composante persistante $\{C_t\}$ et d'une composante non persistante $\{M_t\}$ de la façon suivante :

$$V_t = \sigma^2 C_t M_t,$$

où σ^2 est la variance inconditionnelle.

La composante persistante, prise en compte par le processus $\{C_t\}$, est définie comme

$$C_t = c_0 \prod_{i=1}^N C_t^{(i)},$$

où c_0 est une constante faisant en sorte que $\mathbb{E}[C_t] = 1$. Les sous-composantes $\{C_t^{(i)}\}, i = 1, \dots, N$, partagent une matrice de transition commune \mathbf{P} définie par :

$$\mathbf{P} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix},$$

avec $p \in (0,1)$. [Augustyniak et al. \(2019\)](#) proposent que chacune des sous-composantes $C_t^{(i)}$ du modèle ait son propre espace d'états $E_{C^{(i)}} = \{1, \nu_i\}$, où $\nu_i = (1 - \theta_c) + \theta_c \nu_{i-1}$ avec $\nu_1 > 1$ et $\theta_c \in (0, 1)$. Cette différence des espaces d'états des différentes composantes mène à un espace E_C pour la nouvelle chaîne de Markov satisfaisant

$$E_C = \left\{ \prod_{i \in I} \nu_i, I \subset \{1, 2, \dots, N\} \right\},$$

qui représente l'ensemble de tous les produits d'éléments de l'ensemble $\{\nu_1, \nu_2, \dots, \nu_N\}$ c'est à dire

$$E_C = \left\{ 1, \nu_1, \nu_2, \dots, \nu_N, \nu_1 \nu_2, \dots, \nu_1 \nu_N, \nu_2 \nu_3, \dots, \nu_{N-1} \nu_N, \dots, \prod_{i=1}^N \nu_i \right\}.$$

Cette représentation permet donc d'augmenter significativement le nombre d'états différents de la chaîne de Markov latente du modèle qui est maintenant de 2^N .

La composante non persistante, quant à elle, est prise en compte par un processus $\{M_t\}$ de variables indépendantes et identiquement distribuées selon la fonction de masse suivante :

$$\mathbb{P}(M_t = m_0 \times m_i) = \begin{cases} q(N-1)^{-1}, & \text{si } i = 1, \dots, N-1 \\ 1-q, & \text{si } i = N \end{cases},$$

où $q \in (0, 1)$, $m_1 > 0$, $m_i = (1 - \theta_m) + \theta_m m_{i-1}$ pour $i = 2, \dots, N-1$, $m_N = 1$, $\theta_m \in [0, 1]$ et m_0 est obtenu de sorte à ce que $\mathbb{E}[M_t] = 1$. Cette dernière composante non persistante permet ainsi de capter les changements abrupts qui pourraient être présents dans les données.

Notons que [Augustyniak et al. \(2019\)](#) modélisent non seulement les rendements financiers, utilisant l'équation (1.2.1), mais aussi les variances réalisées, utilisant une équation de type MEM (1.2.4).

En multipliant N matrices, les modèles MSM et CDRS permettent de produire une nouvelle chaîne de Markov comportant $N + 1$ valeurs différentes tandis que le modèle FHMV produit une chaîne de Markov comportant 2^N valeurs différentes.

1.3.3. Un autre HMM à haute dimension : Le modèle DSARV

Les modèles factoriels à chaîne de Markov cachée se basent sur une chaîne de Markov à haute dimension obtenue par le produit de plusieurs chaînes de Markov de faible dimension. Ces modèles ne constituent cependant pas la seule manière élégante de construire une chaîne de Markov à haute dimension. En effet, [Cordis et Kirby \(2014\)](#) proposent le modèle DSARV (*Discrete Stochastic AutoRegressive Volatility*), basé sur une méthode de paramétrisation parcimonieuse d'une matrice $N \times N$. [Cordis et Kirby \(2014\)](#) supposent que le processus latent $\{V_t\}$ est une chaîne de Markov à espace d'états $E_V = \{\nu_1, \nu_2, \dots, \nu_N\}$. Comme ces auteurs l'ont bien précisé, une chaîne de Markov à N états peut être représentée comme un modèle VAR(1) (*Vector AutoRegressive* d'ordre 1). Il suffit de définir \mathbf{x}_t comme étant le vecteur $N \times 1$ indiquant l'état (la j -ième composante de \mathbf{x}_t est égale à 1 et toutes les autres composantes sont nulles lorsque la chaîne de Markov est dans l'état j à la date t), \mathbf{P} comme étant la matrice de transition de la chaîne de Markov et $\boldsymbol{\nu} = (\nu_1 \ \nu_2 \ \dots \ \nu_N)'$. Alors, on a :

$$\begin{aligned} V_t &= \boldsymbol{\nu}' \mathbf{x}_t, \\ \mathbf{x}_{t+1} &= \mathbf{P}' \mathbf{x}_t + \mathbf{e}_{t+1}, \end{aligned}$$

où $\{\mathbf{e}_t\}$ est un processus tel que $\mathbb{E}[\mathbf{e}_t \mid \mathbf{x}_{t-1}, \dots, \mathbf{x}_1] = \mathbf{0}$. [Cordis et Kirby \(2014\)](#) proposent la paramétrisation suivante pour matrice de transition \mathbf{P} :

$$\mathbf{P} = \phi \mathbf{I}_N + (1 - \phi) \mathbf{1}_N \boldsymbol{\pi}',$$

où $\phi \in [0, 1)$, \mathbf{I}_N est la matrice identité d'ordre N , $\mathbf{1}_N$ est le vecteur $N \times 1$ composé uniquement de 1 et $\boldsymbol{\pi}$ représente le vecteur des probabilités stationnaires de la chaîne de Markov $\{V_t\}$, c'est à dire $\mathbb{E}[\mathbf{x}_t] = \boldsymbol{\pi}$. De plus, bien que le modèle DSARV ne soit pas un modèle factoriel puisqu'il n'intègre pas de structure de composantes, il tire son inspiration de deux des modèles factoriels présentés plus haut (MSM, CDRS).

1.4. L'asymétrie dans les rendements financiers : l'effet levier (*leverage effect*)

Tel que précisé à la section 1.1.1, les séries de rendements financiers présentent souvent un effet levier. Ceci signifie qu'un choc négatif significatif sur les rendements financiers au temps t tend à accroître la volatilité des rendements subséquents par un plus grand écart comparativement à un choc positif de la même amplitude. En réalité, les séries de rendements des indices financiers présentent généralement un effet levier, ce qui n'est pas toujours le cas des actions (qui n'en présentent que très peu) ni des taux de change (qui n'en présentent pas du tout). Il est alors important de capter cet effet levier lorsqu'on veut modéliser les indices financiers. Dans la littérature, deux manières principales sont proposées pour prendre en compte cet effet levier dans les modèles à chaîne de Markov cachée.

1.4.1. Effet levier imbriqué dans la matrice de transition

Comme on sait que les valeurs négatives des rendements financiers ont une influence plus grande que les valeurs positives sur les rendements futurs, on peut penser que la matrice de transition des volatilités devrait être influencée différemment selon que les rendements soient positifs ou négatifs. C'est pour cette raison que [Cordis et Kirby \(2014\)](#) proposent d'utiliser une chaîne de Markov non-homogène. En effet, la paramétrisation de la matrice $N \times N$ qu'ils proposent dans leur article peut être généralisée de sorte à obtenir une matrice de transition qui varie dans le temps. Cette approche utilise une fonction logistique des rendements de la manière suivante :

$$\begin{aligned}\pi_{jt} &= \frac{(N-1)!}{(j-1)!(N-j)!} \omega_t^{j-1} (1-\omega_t)^{N-j}, \quad j = 1, \dots, N, \\ \omega_t &= \frac{\exp(\psi + \varphi r_t + \varphi \rho r_{t-1} + \dots + \varphi \rho^{t-1} r_1)}{1 + \exp(\psi + \varphi r_t + \varphi \rho r_{t-1} + \dots + \varphi \rho^{t-1} r_1)}, \\ \mathbf{P}_t &= \phi \mathbf{I}_N + (1 - \phi) \mathbf{1}_N \boldsymbol{\pi}_t',\end{aligned}$$

où $\boldsymbol{\pi}_t = (\pi_{1t} \ \pi_{2t} \ \dots \ \pi_{Nt})'$ et $\phi \in [0,1], \rho \in [0,1]$ et $\psi, \varphi \in \mathbb{R}$ sont des paramètres à estimer. Le signe et la valeur absolue du paramètre φ contrôlent respectivement la direction et la force de l'effet levier. En effet, un signe négatif de φ produit un modèle dans lequel les rendements négatifs sont associés à une augmentation dans la volatilité future espérée. Le paramètre ρ , quant à lui, contrôle le taux auquel l'effet levier diminue dans le temps. Ainsi, si un rendement négatif à la date t signale une augmentation de la volatilité future espérée, alors cette augmentation peut être totalement transitoire ($\rho = 0$), légèrement persistante ($\rho = 0.5$) ou très persistante ($\rho = 0.9$).

1.4.2. Effet levier multiplicatif

Une autre méthode pour considérer l'effet levier est de le prendre directement en compte dans la volatilité du processus. Il suffit, dans ce cas, de multiplier la chaîne de Markov par un processus d'effet de levier. L'intérêt de cette approche est qu'elle permet de déterminer la valeur de l'effet levier à chaque date. [Augustyniak et al. \(2019\)](#) proposent le processus suivant pour tenir compte de l'effet levier :

$$L_t = \prod_{i=1}^{N_L} L_t^{(i)}, \quad \text{avec } L_t^{(i)} = 1 + l_i \frac{|r_{t-i}|}{\sqrt{L_{t-i}}} \mathbf{1}_{\{r_{t-i} < 0\}},$$

où $l_i = \theta_i^{i-1} l_1$, $l_1 > 0$ et $\theta_i \in [0,1]$. Le nombre entier N_L désigne le nombre de rendements passés qui pourraient influencer la volatilité du rendement présent. Le processus L_t proposé par [Augustyniak et al. \(2019\)](#) est un processus prévisible, donc sa valeur à la date t est entièrement déterminée par l'information disponible jusqu'à la date $t-1$. De plus, ce processus peut être interprété comme étant un ensemble d'effets obtenus à partir des rendements passés. Ainsi, on pourra considérer que $L_t^{(i)}$ est actif si $r_{t-i} < 0$ et donc ajouter l'effet $\left(l_i \frac{|r_{t-i}|}{\sqrt{L_{t-i}}} \right)$ entraîné par la sous-composante $L_t^{(i)}$. En d'autres termes, un rendement passé

négligatif ajoute une volatilité supplémentaire d'intensité liée à l'amplitude du rendement et à un paramètre l_i structuré de manière à donner de moins en moins d'importance aux rendements les plus éloignés.

1.5. Discussion

Modéliser les données économiques et financières permet de mieux comprendre l'économie et les marchés financiers afin de prendre les décisions les plus adéquates selon la situation. Les modèles financiers tentent donc de reproduire certaines caractéristiques intéressantes (faits stylisés) spécifiques aux données financières. Les modèles à chaîne de Markov cachée permettent de reproduire la plupart de ces faits stylisés, ce qui explique sans doute le grand nombre d'articles scientifiques portant sur ces modèles financiers au cours des dernières décennies. Malgré leur succès, les modèles à chaîne de Markov cachée classiques présentent quelques désavantages (décroissance rapide de la fonction d'autocorrélation, faible nombre d'états, etc.). Pour pallier à ces inconvénients, une classe de modèles à chaîne de Markov cachée, dite à haute dimension, a vu le jour. Les modèles factoriels à chaîne de Markov cachée (FHMM) sont parcimonieux puisqu'ils permettent de produire un grand nombre de valeurs pour la volatilité avec un nombre de paramètres relativement faible. Ces modèles se basent sur le principe qu'il est possible de créer une chaîne de Markov de haute dimension en multipliant entre elles des chaînes de Markov de dimension plus faible (généralement 2). Par exemple, dans l'un des modèles FHMM (le modèle MSM de [Calvet et Fisher, 2004](#)), la volatilité présente dans les rendements financiers est modélisée en multipliant N chaînes de Markov indépendantes à deux états. Ainsi, une chaîne de Markov à 1024 états peut être obtenue avec seulement quatre paramètres si 10 chaînes sont multipliées. Il faut cependant noter que les 2^N états de la chaîne latente du modèle MSM ne peuvent prendre que $N + 1$ valeurs différentes.

En finance, ces modèles sont proposés pour modéliser soit les rendements financiers ou soit les variances réalisées. Plusieurs études montrent cependant qu'au lieu de modéliser uniquement la série des rendements financiers, le fait d'inclure les variances réalisées permet de mieux capter le niveau de volatilité ([Takahashi et al., 2009](#), [Hansen et al., 2012](#), [Hansen et Huang, 2016](#), [Liu et Maheu, 2018](#), [Takahashi et al., 2021](#)). Ainsi, les modèles joints, qui modélisent simultanément les rendements financiers et les variances réalisées, sont plus performants lorsque l'objectif est de comprendre ou de prédire le comportement des données financières.

Le chapitre 2 formalise le cadre de la modélisation factorielle tout en montrant comment les modèles factoriels proposés dans la littérature tels que les modèles MSM de [Calvet et Fisher \(2004\)](#), CDRS de [Fleming et Kirby \(2013\)](#) et FHMV de [Augustyniak et al. \(2019\)](#) adhèrent à ce cadre. Ce chapitre propose également un nouveau processus, nommé MDSV (*Multifractal Discrete Stochastic Volatility*), qui appartient à la famille des modèles factoriels à chaîne de Markov cachée. Le processus MDSV a une construction particulière lui permettant d'étendre les modèles factoriels existants en autorisant les chaînes de Markov sous-jacentes (composantes) à avoir un espace d'états de cardinalité supérieur à 2. À travers ce chapitre 2, nous étudions les propriétés théoriques et la structure du processus MDSV, puis nous montrons qu'il peut être interprété comme un modèle à volatilité stochastique multicomposantes. Nous indiquons aussi comment ce processus

peut être utilisé pour modéliser les rendements financiers et les variances réalisées, soit séparément ou de manière conjointe.

Chapitre 2

Le modèle MDSV : *Multifractal Discrete Stochastic Volatility*

Dans ce chapitre, nous formalisons le cadre de la modélisation factorielle et démontrons que les modèles factoriels à chaîne de Markov cachée sont étroitement liés aux modèles à volatilité stochastique. Après avoir montré que les modèles à chaîne de Markov cachée à haute dimension présentés au chapitre 1 (le modèle MSM, *Markov Switching Multifractal* de [Calvet et Fisher, 2004](#), le modèle CDRS, *Component-Driven Regime-Switching Volatility* de [Fleming et Kirby, 2013](#) et le modèle FHMV, *Factorial Hidden Markov Volatility* de [Augustyniak et al., 2019](#)) adhèrent à ce cadre, nous introduisons un nouveau processus, nommé MDSV (*Multifractal Discrete Stochastic Volatility*) appartenant à cette classe de modèles. Nous exposons ensuite les propriétés de ce nouveau processus MDSV tout en faisant une analyse théorique. Nous relierons, dans la section 2.3, les rendements financiers et variances réalisées au processus MDSV.

2.1. Cadre général de la modélisation factorielle

Les modèles factoriels à chaîne de Markov ont été introduits par [Ghahramani et Jordan \(1997\)](#). Ils font intervenir une chaîne de Markov à espace d'états de grande dimension obtenue à partir du produit de chaînes de Markov à espace d'états de plus petite dimension. Ces modèles se basent sur le théorème 2.1.1 suivant :

Théorème 2.1.1. *Soit $N \geq 2$ et $\{C_t^{(1)}\}, \{C_t^{(2)}\}, \dots, \{C_t^{(N)}\}$ des chaînes de Markov discrètes, homogènes et indépendantes à espaces d'états finis respectifs $E^{(1)}, E^{(2)}, \dots, E^{(N)}$ et de matrices de transition respectives $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N$.*

Le processus $\{M_t\}$, défini pour chaque t par $M_t = (C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)})$, est une chaîne de Markov discrète et homogène à espace d'états fini $E_M = E^{(1)} \times E^{(2)} \times \dots \times E^{(N)}$ de matrice de transition $\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \dots \otimes \mathbf{P}_N$, où le symbole \otimes représente le produit de Kronecker et \times représente le produit des ensembles.

DÉMONSTRATION. Voir l'annexe A.1. □

Remarque 2.1.1. *Si on peut s'assurer de ne pas agréger les résultats du produit des ensembles $E_M = E^{(1)} \times E^{(2)} \times \dots \times E^{(N)}$ en faisant le produit des composantes de chaque élément, alors le processus $\{C_t\}$ défini pour chaque t par $C_t = \prod_{i=1}^N C_t^{(i)}$ (produit des composantes de M_t) est aussi une chaîne de*

Markov discrète homogène à espace d'états fini $E = E^{(1)} \otimes E^{(2)} \otimes \dots \otimes E^{(N)}$ et de matrice de transition $\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \dots \otimes \mathbf{P}_N$.

L'exemple 2.1.1 suivant permet de mieux comprendre la remarque 2.1.1.

Exemple 2.1.1. Soit $\{C_t^{(1)}\}$ et $\{C_t^{(2)}\}$ deux chaînes de Markov discrètes, homogènes et indépendantes à espaces d'états finis respectifs $E^{(1)} = \{1, 2\}$ et $E^{(2)} = \{1, 2\}$ et de matrices de transition respectives \mathbf{P}_1 et \mathbf{P}_2 définies par :

$$\mathbf{P}_1 = \begin{pmatrix} p_1 & 1 - p_1 \\ p_2 & 1 - p_2 \end{pmatrix}, \quad \mathbf{P}_2 = \begin{pmatrix} q_1 & 1 - q_1 \\ q_2 & 1 - q_2 \end{pmatrix}.$$

Le processus $\{C_t\}$ défini à toute date t par $C_t = C_t^{(1)} C_t^{(2)}$ qui prend ses valeurs dans l'espace

$$E = E^{(1)} \otimes E^{(2)} = \{1, 2\} \otimes \{1, 2\} = \{1, 2, 2, 4\},$$

est une chaîne de Markov homogène de matrice de transition

$$\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2 = \begin{pmatrix} p_1 q_1 & p_1 (1 - q_1) & (1 - p_1) q_1 & (1 - p_1) (1 - q_1) \\ p_1 q_2 & p_1 (1 - q_2) & (1 - p_1) q_2 & (1 - p_1) (1 - q_2) \\ p_2 q_1 & p_2 (1 - q_1) & (1 - p_2) q_1 & (1 - p_2) (1 - q_1) \\ p_2 q_2 & p_2 (1 - q_2) & (1 - p_2) q_2 & (1 - p_2) (1 - q_2) \end{pmatrix}.$$

Notons qu'ici, nous considérons que l'état de valeur 2 formé par $C_t^{(1)} = 1$ et $C_t^{(2)} = 2$ est différent de l'état de même valeur 2 qui est formé par $C_t^{(1)} = 2$ et $C_t^{(2)} = 1$. En effet, le processus à trois états $\{1, 2, 4\}$, obtenu en agrégeant ces deux valeurs n'est généralement pas une chaîne de Markov.

2.1.1. Présentation du cadre général

Dans le cadre de cette thèse, on dira qu'un modèle à chaîne de Markov cachée est un modèle factoriel lorsque la chaîne de Markov $V = \{V_t\}$ de ce modèle est construite à partir du produit de N chaînes de Markov indépendantes, notées $C^{(i)} = \{C_t^{(i)}\}, i = 1, 2, \dots, N$:

$$V_t = \sigma^2 \frac{\prod_{i=1}^N C_t^{(i)}}{\prod_{i=1}^N \mathbb{E}[C_t^{(i)}]}, \quad (2.1.1)$$

où $\sigma^2 \in (0, \infty)$ est un paramètre correspondant à la variance inconditionnelle, c'est-à-dire $\mathbb{E}[V_t] = \sigma^2$. On suppose également que chaque chaîne de Markov $C^{(i)}, i = 1, 2, \dots, N$, qu'on désignera par « composante », évolue dans un espace d'états, noté $\nu^{(i)}$, de cardinalité K et que sa matrice de transition est modélisée par :

$$\mathbf{P}^{(i)} = \phi^{(i)} \mathbf{I}_K + (1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'}, \quad (2.1.2)$$

où \mathbf{I}_K est la matrice identité de dimension $K \times K$, $\mathbf{1}_K$ est le vecteur $K \times 1$ composé uniquement de 1, $\phi^{(i)} \in [0, 1)$ est un paramètre, et $\boldsymbol{\pi}^{(i)}$ est un vecteur paramètre de probabilités qui correspond à la distribution stationnaire de $C^{(i)}$, donc $\mathbb{E}[C_t^{(i)}] = \boldsymbol{\pi}^{(i)' \nu^{(i)}}$. En accord avec le théorème 2.1.1 et la remarque 2.1.1, le

processus V est une chaîne de Markov puisqu'il est constitué d'un produit de chaînes de Markov. De plus, il est caractérisé par son espace d'états de cardinalité K^N , noté $\boldsymbol{\nu}$, donné par :

$$\boldsymbol{\nu} = \frac{\sigma^2}{\prod_{i=1}^N [\boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)}]} \left[\otimes_{i=1}^N \left(\boldsymbol{\nu}^{(i)} \right) \right],$$

sa matrice de transition, notée \mathbf{P} , donnée par :

$$\mathbf{P} = \otimes_{i=1}^N \left(\mathbf{P}^{(i)} \right),$$

et sa distribution stationnaire, notée $\boldsymbol{\pi}$, donnée par :

$$\boldsymbol{\pi} = \otimes_{i=1}^N \left(\boldsymbol{\pi}^{(i)} \right).$$

Dans cette construction, la structure particulière des matrices de transitions $\mathbf{P}^{(i)}$ (équation (2.1.2)) implique que le logarithme de $C^{(i)}$ admet une représentation autorégressive d'ordre 1 (AR(1)) telle qu'indiquée dans la proposition 2.1.1 suivante :

Proposition 2.1.1. *Soit V , la chaîne de Markov définie par l'équation (2.1.1). En posant $h_t^{(i)} = \log \left(C_t^{(i)} \right)$, on obtient que :*

(i) *Le processus $\{h_t^{(i)}\}$ a une représentation autorégressive d'ordre 1 :*

$$h_{t+1}^{(i)} = c_i + \phi^{(i)} h_t^{(i)} + u_{t+1}^{(i)}, \quad (2.1.3)$$

où $c_i = (1 - \phi^{(i)}) \boldsymbol{\pi}^{(i)'} \log \boldsymbol{\nu}^{(i)}$ et $u_t^{(i)}$ est un processus vérifiant $\mathbb{E} \left[u_t^{(i)} \mid h_{t-1}^{(i)}, \dots, h_1^{(i)} \right] = 0$.

(ii) *Le processus V vérifie la relation*

$$V_t = \frac{\sigma^2}{\prod_{i=1}^N [\boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)}]} \exp \left(\sum_{i=1}^N h_t^{(i)} \right). \quad (2.1.4)$$

DÉMONSTRATION. Voir l'annexe A.2. □

L'équation (2.1.4) implique que notre modèle factoriel à chaîne de Markov cachée peut être interprété comme un modèle à volatilité stochastique multicomposantes, tel que présenté à la section 1.2.2, et peut être vu comme une version discrète du modèle proposé par Andersen et Bollerslev (1997). En effet, la classe des modèles à volatilité stochastique est l'une des plus populaires dans la modélisation financière. Les modèles à volatilité stochastique, aussi bien discrets que continus, ont été très utilisés en finance théorique, empirique et en économétrie financière (ex : Hull et White, 1987, Duffie et Singleton, 1990, Heston, 1993, Bates, 1996, Alizadeh *et al.*, 2002, Takahashi *et al.*, 2021).

Ce cadre général de la modélisation factorielle englobe les modèles factoriels que nous avons mentionnés au chapitre 1 dans la mesure où la chaîne de Markov cachée dans ces modèles est construite par une structure multiplicative. De plus, pour $K = 2$, la matrice de transition $\mathbf{P}^{(i)}$ définie par l'équation (2.1.2) est arbitraire et basée sur deux paramètres. Ceci fait qu'en imposant des contraintes adéquates sur ces paramètres, on peut facilement montrer que le modèle MSM de Calvet et Fisher (2004), CDRS de Fleming et Kirby (2013) et le modèle FHMV de Augustyniak *et al.* (2019) correspondent à cette définition de modèles factoriels.

2.1.2. Quelques propriétés théoriques des modèles factoriels

Dans cette section, nous présentons quelques propriétés théoriques du modèle factoriel à chaîne de Markov cachée que nous avons introduit à la section 2.1.1. La proposition 2.1.2 présente des propriétés de la matrice de transition \mathbf{P} de la chaîne de Markov V .

Proposition 2.1.2 (Propriétés de la matrice de transition \mathbf{P}). *La matrice de transition $\mathbf{P} = \otimes_{i=1}^N (\mathbf{P}^{(i)})$ de V satisfait les propriétés suivantes :*

(i) *La matrice de transition $\mathbf{P}^{(i)}$ de $C^{(i)}$ a une unique valeur propre non unitaire égale à $\phi^{(i)}$. La multiplicité algébrique de cette valeur propre est $K - 1$.*

(ii) *La matrice de transition \mathbf{P} de V a 2^N valeurs propres données par :*

$$\left\{ \prod_{i \in I} \phi^{(i)} : I \in A \right\},$$

où A désigne l'ensemble des parties de $\{1, 2, \dots, N\}$. La multiplicité algébrique de la valeur propre $\prod_{i \in I} \phi^{(i)}$, $I \in A$ est $(K - 1)^{|I|}$.

(iii) *La matrice de transition en k pas de temps \mathbf{P}^k de V admet la représentation suivante :*

$$\mathbf{P}^k = \otimes_{i=1}^N \left\{ \left(\phi^{(i)} \right)^k \mathbf{I}_K + \left[1 - \left(\phi^{(i)} \right)^k \right] \mathbf{1}_K \boldsymbol{\pi}^{(i)'} \right\},$$

et converge vers $\boldsymbol{\Pi} = \mathbf{1}_{K^N} \boldsymbol{\pi}'$ lorsque $k \rightarrow \infty$ avec un taux de convergence asymptotique donné par :

$$\mathbf{P}^k - \boldsymbol{\Pi} = \mathcal{O}(k^{K-2} a^k), \quad (2.1.5)$$

et un taux de convergence non-asymptotique donné par :

$$\|\mathbf{P}^k - \boldsymbol{\Pi}\|_\infty \leq 2 - 2(1 - a^k)^N, \quad (2.1.6)$$

où $a = \max_{i=1, \dots, N} \phi^{(i)}$ et $\|\cdot\|_\infty$ désigne la norme maximum des sommes des valeurs absolues des lignes.

DÉMONSTRATION. Voir l'annexe A.3. □

La proposition 2.1.2 montre que le nombre N de composantes, d'une part, et le nombre K d'états des composantes, d'autre part, influencent la vitesse de convergence de \mathbf{P}^k lorsque k augmente. Par exemple, lorsque $N = 1$ et $K = 2$, on trouve $\|\mathbf{P}^k - \boldsymbol{\Pi}\|_\infty \leq 2a^k$ et $\mathbf{P}^k - \boldsymbol{\Pi} = \mathcal{O}(a^k)$. En effet, les équations (2.1.6) et (2.1.5) indiquent que les grandes valeurs de N et de K entraînent une plus faible vitesse de convergence de \mathbf{P}^k lorsque $k = 1, 2, \dots$.

En se basant sur la théorie des chaînes de Markov standard (voir Hamilton, 1994, chapitre 22), le niveau de persistance dans un modèle à chaîne de Markov cachée peut être caractérisé par la vitesse à laquelle la fonction d'autocovariance $\text{Cov}(V_{t+k}, V_t)$ de la chaîne de Markov V se rapproche de 0 lorsque k augmente. Hamilton (1994) démontre que cette fonction d'autocovariance vérifie, pour tout $k = 1, 2, \dots$, la relation

$$\text{Cov}(V_{t+k}, V_t) = \boldsymbol{\pi}' \boldsymbol{\Upsilon} \mathbf{P}^k \boldsymbol{\nu} - (\boldsymbol{\pi}' \boldsymbol{\nu})^2 = \boldsymbol{\pi}' \boldsymbol{\Upsilon} (\mathbf{P}^k - \boldsymbol{\Pi}) \boldsymbol{\nu}, \quad (2.1.7)$$

où Υ désigne la matrice diagonale contenant ν c'est-à-dire la matrice vérifiant $\Upsilon \mathbf{1}_{K^N} = \nu$. Cette relation confirme que la fonction d'autocovariance converge vers 0 lorsque $k \rightarrow \infty$ puisque \mathbf{P}^k converge vers $\mathbf{\Pi}$ lorsque $k \rightarrow \infty$. De plus, cette relation indique que la vitesse à laquelle la fonction d'autocovariance converge vers 0 dépend du taux de convergence de la matrice de transition \mathbf{P}^k vers sa limite $\mathbf{\Pi}$, qui est $\mathcal{O}(k^{K-2}a^k)$ d'après la proposition 2.1.2. Par conséquent, la décroissance de la fonction d'autocovariance de V est plus lente à mesure que le nombre d'états, K , des composantes $C^{(i)}$ augmente et à mesure que la persistance la plus élevée a augmente.

Le théorème 2.1.2 présente la distribution du temps de séjour de la chaîne de Markov V dans tout sous-ensemble ν_s de ν .

Théorème 2.1.2 (Temps de séjour de la chaîne V). *Soit $V = \{V_t\}$, une chaîne de Markov homogène, irréductible, de matrice de transition \mathbf{P} et d'espace d'états ν . Désignons par π la distribution stationnaire de V et soit ν_s , un sous-espace de ν . On peut réorganiser ν , \mathbf{P} et π de sorte que la partition $\nu = \left(\nu'_s \quad \overline{\nu}'_s \right)'$ induise la décomposition de \mathbf{P} en quatre sous-matrices et la décomposition de π en deux sous-vecteurs :*

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{\nu_s} & \mathbf{P}_{\nu_s \overline{\nu}_s} \\ \mathbf{P}_{\overline{\nu}_s \nu_s} & \mathbf{P}_{\overline{\nu}_s} \end{pmatrix}, \quad \pi = \left(\pi'_s \quad \overline{\pi}'_s \right)'.$$

Le temps de séjour de la chaîne V dans le sous-espace ν_s , noté N_{ν_s} , c'est-à-dire le nombre de périodes successives que la chaîne V passe dans le sous-espace ν_s est distribué selon :

$$\mathbb{P}(N_{\nu_s} = k) = \frac{\pi'_s (\mathbf{I}_{n_s} - \mathbf{P}_{\nu_s}) \mathbf{P}_{\nu_s}^{k-1} (\mathbf{I}_{n_s} - \mathbf{P}_{\nu_s}) \mathbf{1}'_{n_s}}{\pi'_s (\mathbf{I}_{n_s} - \mathbf{P}_{\nu_s}) \mathbf{1}'_{n_s}}, \quad k = 1, 2, \dots, \quad (2.1.8)$$

où $n_s = |\nu_s|$, \mathbf{I}_{n_s} représente la matrice identité de dimension $n_s \times n_s$ et $\mathbf{1}_{n_s}$ est le vecteur $n_s \times 1$ composé uniquement de 1.

DÉMONSTRATION. Le résultat s'obtient directement du théorème 2.6 et du corollaire 2.7 de Rubino et Sericola (1989). □

Dans les modèles à chaîne de Markov cachée, le temps de séjour de la chaîne de Markov dans ses états est une notion importante. En effet, Rydén *et al.* (1998) expliquent que les modèles à chaîne de Markov cachée ne peuvent pas reproduire la lente décroissance de la fonction d'autocorrélation du carré des rendements financiers empiriquement observée. Ce manque de flexibilité des modèles à chaîne de Markov cachée pour modéliser la dépendance temporelle d'ordre supérieur peut s'expliquer par le temps de séjour dans les états cachés distribué géométriquement (voir remarque 1.3.1). Le théorème 2.1.2 nous indique qu'il est possible de produire un temps de séjour non géométriquement distribué avec la famille des modèles factoriels à chaîne de Markov cachée dont le cadre est présenté dans cette section. Il suffit de spécifier la chaîne V de sorte qu'elle ait des états qui se répètent¹. Ce faisant, on crée un modèle présentant des propriétés similaires à celles d'un modèle à chaîne semi-markovienne cachée. Les modèles à chaîne semi-markovienne cachée sont une généralisation des modèles à chaîne de Markov cachée qui autorisent des distributions plus flexibles du

1. Plus de détails sur cette spécification sont donnés à la section 2.2 suivante

temps de séjour dans les états visités. Bulla et Bulla (2006) montrent que le fait d'utiliser un modèle à chaîne semi-markovienne cachée permet de mieux reproduire la lente décroissance de la fonction d'autocorrélation du carré des rendements financiers grâce à la flexibilité obtenue par la distribution du temps de séjour.

2.2. Le processus MDSV

Les modèles factoriels à chaîne de Markov cachée ont l'avantage de la parcimonie par rapport aux modèles à chaîne de Markov cachée classiques. Ils permettent de produire un grand nombre d'états de la chaîne de Markov avec un nombre de paramètres relativement faible. Cependant, les modèles factoriels à chaîne de Markov cachée de la littérature se basent sur une hypothèse intrinsèque contraignante stipulant que "chacune des composantes n'a que deux états". Le cadre de modélisation factorielle présenté à la section 2.1 permet de lever cette contrainte. Ce cadre reste néanmoins trop général pour constituer un modèle avec un nombre raisonnable de paramètres. Dans cette section, nous proposons le processus MDSV (*Multifractal Discrete Stochastic Volatility*), qu'on utilise pour construire un modèle factoriel à chaîne de Markov cachée qui autorise ses différentes composantes à avoir plus de deux états. Il s'agit d'imposer les contraintes adéquates aux vecteurs de paramètres $\nu^{(i)}$, $\pi^{(i)}$ et aux paramètres $\phi^{(i)}$, $i = 1, 2, \dots, N$ dans le cadre de modélisation factoriel afin d'avoir un modèle parcimonieux et adapté aux données financières.

2.2.1. Présentation du processus

Le processus MDSV est conçu pour modéliser un facteur de risque à valeurs positives dans un modèle économétrique, telle que la volatilité. Il correspond au processus $V = \{V_t\}$ défini pour chaque t par :

$$V_t = \sigma^2 \frac{\prod_{i=1}^N C_t^{(i)}}{\prod_{i=1}^N \mathbb{E} [C_t^{(i)}]}, \quad (2.2.1)$$

où $\sigma^2 > 0$ est une constante et les composantes $C_t^{(i)} = \{C_t^{(i)}\}$, $i = 1, 2, \dots, N$ sont des chaînes de Markov d'espace d'états $\nu^{(i)}$, de distribution stationnaire $\pi^{(i)}$ et dont la matrice de transition vérifie

$$\mathbf{P}^{(i)} = \phi^{(i)} \mathbf{I}_K + (1 - \phi^{(i)}) \mathbf{1}_K \pi^{(i)'}. \quad (2.2.2)$$

On suppose cependant que les niveaux de persistance $\phi^{(i)}$, $i = 1, 2, \dots, N$, sont structurés dans un ordre décroissant à travers la relation :

$$\phi^{(i)} = a^{b^{i-1}}, \quad i = 1, \dots, N, \quad (2.2.3)$$

où $a \in (0, 1)$ et $b \in [1, \infty)$. Cette spécification est introduite par Calvet et Fisher (2001) en relation avec la discrétisation d'un processus de Poisson et est utilisée par Calvet et Fisher (2004) dans la définition du modèle MSM. Cette relation entraîne une décroissance exponentielle des persistances puisque $\phi^{(i+1)} = (\phi^{(i)})^b$ et donc $\phi^{(1)} > \phi^{(2)} > \dots > \phi^{(N)}$. On crée ainsi une structure hiérarchique dans la chaîne de Markov V .

Pour créer un modèle parcimonieux, on suppose également que toutes les composantes $C^{(i)}, i = 1, 2, \dots, N$, partagent le même espace d'états et la même distribution stationnaire, qui sont :

$$\boldsymbol{\nu}^{(i)} = \boldsymbol{\nu}^{(1)} = \left(\nu_1^{(1)} \quad \nu_2^{(1)} \quad \dots \quad \nu_K^{(1)} \right)', \quad i = 1, \dots, N, \quad (2.2.4)$$

$$\boldsymbol{\pi}^{(i)} = \boldsymbol{\pi}^{(1)} = \left(\pi_1^{(1)} \quad \pi_2^{(1)} \quad \dots \quad \pi_K^{(1)} \right), \quad i = 1, \dots, N. \quad (2.2.5)$$

Les éléments $\nu_j^{(1)}, j = 1, 2, \dots, K$ de l'espace d'états sont paramétrisés comme une fonction d'un paramètre unique $\nu_0 \in (0,1)$ en posant :

$$\nu_j^{(1)} = \nu_0 \left(\frac{2 - \nu_0}{\nu_0} \right)^{j-1}, \quad j = 1, \dots, K. \quad (2.2.6)$$

Cette paramétrisation est équivalente à une structure log-linéaire des états puisque

$$\log \nu_j^{(1)} = \log \nu_0 + (j - 1) \log \left(\frac{2 - \nu_0}{\nu_0} \right), \quad j = 1, 2, \dots, K$$

où $\frac{2 - \nu_0}{\nu_0} > 1$. Par conséquent, nous avons $\nu_1^{(1)} < \nu_2^{(1)} < \dots < \nu_K^{(1)}$.

On suppose aussi que les probabilités stationnaires $\pi_j^{(1)}, j = 1, 2, \dots, K$, présentent une structure binomiale de paramètre $\omega \in (0,1)$:

$$\pi_j^{(1)} = \binom{K-1}{j-1} \omega^{j-1} (1-\omega)^{K-j}, \quad j = 1, \dots, K. \quad (2.2.7)$$

Remarquons que la distribution inconditionnelle de $\log C_t^{(i)}$ implique deux paramètres, ν_0 et ω , et peut être représentée comme une transformation d'une distribution binomiale. En effet, si X désigne une variable distribuée selon une loi binomiale de support $\{0, 1, \dots, K-1\}$ et de probabilité de succès ω alors la distribution inconditionnelle de $\log C_t^{(i)}$ correspond à celle de $\log \nu_0 + X \log \left(\frac{2 - \nu_0}{\nu_0} \right)$. Cela implique que pour K suffisamment grand, la distribution inconditionnelle de $C_t^{(i)}$ est approximativement log-normale.

Au final, nous désignons par $\text{MDSV}(N, K)$ le processus V construit à l'aide du produit de N chaînes de Markov $C^{(i)}, i = 1, \dots, N$, de dimension K . Le processus V est ainsi une chaîne de Markov de dimension K^N avec un espace états, noté $\boldsymbol{\nu}$, donné par ² :

$$\boldsymbol{\nu} = \frac{\sigma^2}{[\boldsymbol{\pi}^{(1)} \boldsymbol{\nu}^{(1)}]^N} \left(\boldsymbol{\nu}^{(1)} \right)^{\otimes N},$$

une matrice de transition, notée \mathbf{P} , donnée par :

$$\mathbf{P} = \otimes_{i=1}^N \left(\mathbf{P}^{(i)} \right),$$

et une distribution stationnaire, notée $\boldsymbol{\pi}$, donnée par :

$$\boldsymbol{\pi} = \left(\boldsymbol{\pi}^{(1)} \right)^{\otimes N}.$$

Le processus MDSV dépend ainsi de cinq paramètres, $\sigma^2, \omega, \nu_0, a$ et b , peu importe le choix des dimensions N et K .

2. L'exposant $\otimes N$ représente la N -ième puissance de Kronecker (ex : la k -ième puissance de Kronecker est définie par induction pour $k \in \mathbb{N}$ par $(\boldsymbol{\nu}^{(1)})^{\otimes 1} = (\boldsymbol{\nu}^{(1)})$ et $(\boldsymbol{\nu}^{(1)})^{\otimes k} = (\boldsymbol{\nu}^{(1)}) \otimes (\boldsymbol{\nu}^{(1)})^{\otimes (k-1)}, k = 2, 3, \dots$).

Cette construction spécifique du processus MDSV fait en sorte que son espace d'états ν contient $N(K - 1) + 1$ valeurs distinctes bien qu'il ait une cardinalité de K^N (c'est-à-dire qu'il y a des valeurs répétées dans l'espace d'états). Un fait intéressant est que le processus V n'est pas une chaîne de Markov sur ces valeurs distinctes et par conséquent le temps passé dans un niveau de volatilité donné (non extrême) n'est pas distribué selon une loi géométrique. Ceci constitue une caractéristique importante de notre modèle et explique pourquoi le processus MDSV a le potentiel de générer un niveau de persistance dans la volatilité qui est plus élevé que les chaînes de Markov conventionnelles basées sur un faible nombre d'états non contraints. Les propriétés théoriques du processus MDSV sont explorées plus en détail dans la section 2.2.3.

2.2.2. Interprétations et lien avec la littérature

La construction du processus MDSV permet de généraliser plusieurs modèles de volatilité à chaîne de Markov cachée proposés dans la littérature tels que le modèle MSM de [Calvet et Fisher \(2004\)](#), le modèle CDRS de [Fleming et Kirby \(2013\)](#), le modèle DSVARV de [Cordis et Kirby \(2014\)](#) et le modèle FHMV de [Augustyniak et al. \(2019\)](#). Par exemple, la spécification log-linéaire du processus DSARV de [Cordis et Kirby \(2014\)](#) coïncide avec le processus MDSV si on suppose qu'il n'existe qu'une seule composante ($N = 1$). Puisque le processus DSARV est obtenu avec une unique chaîne de Markov dont la matrice de transition présente la structure en (2.2.2), le processus MDSV peut être interprété comme un modèle multicomposantes DSARV.

Comme le processus MDSV, les modèles MSM, CDRS et FHMV sont des modèles factoriels qui respectent le cadre de la modélisation factorielle présenté à la section 2.1.1. Plus particulièrement, le modèle binomial MSM de [Calvet et Fisher \(2004\)](#) est un cas particulier du modèle MDSV en fixant $\omega = 1/2$, $K = 2$ et en imposant que les fréquences de chaque composante $C^{(i)}$ soient $\gamma_i = 2 \left(1 - \phi^{(i)}\right)^3$. De même, le processus CDRS($N, N, N - 1$) de [Fleming et Kirby \(2013\)](#) est très semblable au processus MDSV($N, 2$). Ces deux processus diffèrent uniquement par la paramétrisation du vecteur de volatilités et des paramètres $\phi^{(i)}, i = 1, 2, \dots, N$. Le modèle CDRS laisse les paramètres $\phi^{(i)}, i = 1, 2, \dots, N$ libres, alors que le modèle MDSV imposent la relation 2.2.3. Aussi, en imposant $\beta_0 = \log(\nu_0)$ et $\beta_1 = \log\left(\frac{2 - \nu_0}{\nu_0}\right)$ dans la paramétrisation du vecteur de volatilités dans le modèle CDRS, on obtient celle du modèle MDSV⁴. Le modèle MDSV($N, 2$) apparaît ainsi plus parcimonieux que le modèle CDRS. Par ailleurs, en fixant $b = 1$ et en autorisant chaque chaîne à avoir un espace d'états spécifique $\left\{\nu_0^{(i)}, 2 - \nu_0^{(i)}\right\}$ dans le processus MDSV($N, 2$), on obtient un processus très similaire au modèle FHMV de [Augustyniak et al. \(2019\)](#). Cela implique que les modèles MSM, CDRS et FHMV peuvent également être interprétés comme des modèles à volatilité stochastique multicomposantes. Cependant, notre processus MDSV généralise ces spécifications pour permettre aux composantes d'avoir un espace d'états de dimension K supérieure à deux. La possibilité de considérer $K > 2$ se traduit par une plus grande flexibilité pour modéliser les états de volatilité et la persistance. De

3. voir la section 1.3.2.1 pour la définition du modèle MSM.

4. voir la section 1.3.2.2 pour la définition du modèle CDRS.

plus, cette flexibilité peut permettre d'obtenir des fonctions d'autocorrélation de V suffisamment lentes pour répliquer celles observées empiriquement.

La proposition 2.1.1 implique que le processus MDSV peut être relié aux modèles à volatilité stochastique multicomposantes proposé par Andersen et Bollerslev (1997). Leur modèle s'appuie sur l'idée que la volatilité globale du marché représente la manifestation de nombreux processus d'informations hétérogènes ; certains avec des dépendances de volatilité à court terme, d'autres possédant des dépendances de volatilité plus persistantes. Andersen et Bollerslev (1997) modélisent ainsi la log-volatilité du marché comme une agrégation de processus AR(1), similairement à $\log C_t^{(i)}$ à l'équation (2.1.3). Le design du processus MDSV associe, à chaque composante $C^{(i)}$, un unique niveau de persistance. Chaque chaîne de Markov $C^{(i)}$ peut donc être vue comme une composante de l'économie ayant une fréquence spécifique liée à la persistance $\phi^{(i)}$. Les composantes peuvent aussi être interprétées comme des nouvelles de différents niveaux de persistance arrivant sur le marché financier et influençant la volatilité proportionnellement à l'importance de la nouvelle. Le modèle MDSV peut ainsi être vu comme une version discrète du modèle proposé par Andersen et Bollerslev (1997) et hérite ainsi de l'élégante interprétation de leur modèle.

2.2.3. Analyses théoriques et propriétés

La proposition 2.2.1 donne les fonctions d'autocovariance et d'autocorrélation du processus V défini à l'équation (2.2.1).

Proposition 2.2.1 (Structure d'autocovariance du processus MDSV). *Soit φ un nombre réel et les deux quantités suivantes :*

$$\psi_\varphi = \frac{\omega(2 - \nu_0)^{2\varphi} + (1 - \omega)\nu_0^{2\varphi}}{[\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi]^2}, \quad \text{et} \quad \tau_\varphi = \frac{\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi}{[\omega(2 - \nu_0) + (1 - \omega)\nu_0]^\varphi}.$$

Pour $k \in \mathbb{N}$ et $k \geq 1$, on a :

$$(i) \quad \mathbb{E}[V_t^\varphi] = \sigma^{2\varphi} \tau_\varphi^{N(K-1)}. \quad (2.2.8)$$

$$(ii) \quad \text{Var}(V_t^\varphi) = \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \left(\psi_\varphi^{N(K-1)} - 1 \right). \quad (2.2.9)$$

$$(iii) \quad \text{Cov}(V_t^\varphi, V_{t+k}^\varphi) = \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \left(\prod_{i=1}^N \left(1 + (\psi_\varphi^{K-1} - 1) \phi^{(i)k} \right) - 1 \right). \quad (2.2.10)$$

$$(iv) \quad \text{Corr}(V_t^\varphi, V_{t+k}^\varphi) = \frac{\prod_{i=1}^N \left(1 + (\psi_\varphi^{K-1} - 1) \phi^{(i)k} \right) - 1}{\psi_\varphi^{N(K-1)} - 1}. \quad (2.2.11)$$

En particulier pour $\varphi = 1$, on a :

$$\psi_1 = \frac{\nu_0^2 + 4\omega(1 - \nu_0)}{(\nu_0 + 2\omega(1 - \nu_0))^2}, \quad \text{et} \quad \tau_1 = 1.$$

DÉMONSTRATION. Voir l'annexe A.4. □

A partir de la proposition 2.2.1, on remarque que la fonction d'autocovariance du processus MDSV est une fonction croissante de N et K . En effet, si f_N désigne la fonction d'autocovariance $f_N(k) = \mathbb{C}ov(V_t, V_{t+k})$, on remarque facilement que

$$\begin{aligned} \prod_{i=1}^N \left(1 + (\psi_1^{K-1} - 1) \phi^{(i)k}\right) &= \left(1 + (\psi_1^{K-1} - 1) \phi^{(N)k}\right) \left[\prod_{i=1}^{N-1} \left(1 + (\psi_1^{K-1} - 1) \phi^{(i)k}\right) \right], \\ &\geq \prod_{i=1}^{N-1} \left(1 + (\psi_1^{K-1} - 1) \phi^{(i)k}\right), \end{aligned}$$

puisque $(\psi_1^{K-1} - 1) \phi^{(N)k} \geq 0$. Ainsi,

$$f_N(k) = \sigma^4 \left(\prod_{i=1}^N \left(1 + (\psi_1^{K-1} - 1) \phi^{(i)k}\right) - 1 \right) \geq \sigma^4 \left(\prod_{i=1}^{N-1} \left(1 + (\psi_1^{K-1} - 1) \phi^{(i)k}\right) - 1 \right) = f_{N-1}(k).$$

De manière analogue, puisque $\psi_1 > 1$, si on pose $g_K(k) = \mathbb{C}ov(V_t, V_{t+k})$, on montre que

$$g_K(k) = \sigma^4 \left(\prod_{i=1}^N \left(1 + (\psi_1^{K-1} - 1) \phi^{(i)k}\right) - 1 \right) > \sigma^4 \left(\prod_{i=1}^N \left(1 + (\psi_1^{K-2} - 1) \phi^{(i)k}\right) - 1 \right) = g_{K-1}(k).$$

Comme la fonction d'autocovariance, la fonction d'autocorrélation dépend du nombre de composantes N et du nombre d'états K des composantes. Ces deux hyperparamètres (N et K) offrent une grande flexibilité à la fonction d'autocorrélation et permettent ainsi de refléter différentes allures. Le processus MDSV est donc adapté pour modéliser des données présentant une fonction d'autocorrélation qui décroît lentement.

Proposition 2.2.2 (Distribution stationnaire et distribution du temps de séjour). *L'espace d'états ν de dimension K^N du processus V défini à l'équation (2.2.1) contient $N(K-1) + 1$ valeurs distinctes, notées dans un ordre croissant par $x_1, x_2, \dots, x_{N(K-1)+1}$ où :*

$$x_k = x_0 \nu_0^{N-k+1} (2 - \nu_0)^{k-1}, \quad k = 1, 2, \dots, N(K-1) + 1,$$

$$\text{avec } x_0 = \frac{\sigma^2}{[\boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)}]^N} = \frac{\sigma^2}{[\nu_0^{2-K} (\omega(2 - \nu_0) + (1 - \omega)\nu_0)^{K-1}]^N}.$$

(i) *Le nombre n_k de fois que la valeur x_k est répétée dans l'espace d'états ν est*

$$n_k = \sum_{m=0}^{\lfloor \frac{k-1}{K} \rfloor} (-1)^m \binom{N+k-mK-2}{N-1} \binom{N}{m},$$

où le symbole $\lfloor \cdot \rfloor$ représente la partie entière⁵.

(ii) *La distribution stationnaire du processus V sur $x_1, x_2, \dots, x_{N(K-1)+1}$ présente une structure binomiale :*

$$\mathbb{P}[V_t = x_k] = \binom{N(K-1)}{k-1} \omega^{k-1} (1 - \omega)^{N(K-1)+1-k}, \quad k = 1, 2, \dots, N(K-1) + 1. \quad (2.2.12)$$

5. $\lfloor x \rfloor$ est l'unique nombre entier vérifiant $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

(iii) La distribution du temps de séjour du processus V dans le sous-ensemble des états ayant pour valeur x_k est donnée par :

$$\mathbb{P}(N_k = n) = \frac{\mathbf{1}'_{n_k} (\mathbf{I}_{n_k} - \mathbf{P}_{x_k}) \mathbf{P}_{x_k}^{n-1} (\mathbf{I}_{n_k} - \mathbf{P}_{x_k}) \mathbf{1}_{n_k}}{\mathbf{1}'_{n_k} (\mathbf{I}_{n_k} - \mathbf{P}_{x_k}) \mathbf{1}_{n_k}}, \quad n = 1, 2, \dots, \quad (2.2.13)$$

où N_k désigne la variable aléatoire qui compte le nombre de périodes successives pendant lesquelles la valeur prise par le processus V est x_k et \mathbf{P}_{x_k} représente la sous-matrice de \mathbf{P} qui contient les probabilités de transitions dans le sous-espace d'états de valeur x_k .

DÉMONSTRATION. Voir l'annexe A.5. □

L'équation (2.2.13) nous indique que si $n_k > 1$, la distribution du temps de séjour du processus V dans l'état x_k n'est pas géométrique. Puisque $n_k > 1$ pour $k = 2, \dots, N(K-1)$, alors le temps passé dans toutes les valeurs d'états, sauf les valeurs extrêmes, n'est pas distribué géométriquement. De plus, remarquons que cette distribution dépend de n_j et peut donc être différente d'une valeur de l'espace d'états à une autre. Cet effet semi-markovien qui est créé produit alors une décroissance plus lente de la fonction d'autocorrélation (Bulla et Bulla, 2006, Maruotti *et al.*, 2019). Par ailleurs, la distribution stationnaire de $\log V_t$ sur $\log x_1, \log x_2, \dots, \log x_{N(K-1)+1}$ peut être représentée comme une transformation linéaire d'une distribution binomiale. Cela implique que pour $N(K-1)+1$ suffisamment grand, la distribution inconditionnelle de V_t est approximativement log-normale. Notre hypothèse est donc conforme à la littérature de la volatilité réalisée qui suggère que la distribution marginale de la volatilité est proche d'une log-normale (voir, par exemple, Andersen *et al.*, 2001).

2.3. Modèles à chaîne de Markov cachée basés sur le processus MDSV

Dans cette section, nous lions le processus MDSV aux données puis présentons les méthodes d'estimation et de prévision du modèle ainsi obtenu.

2.3.1. Liaison du processus MDSV aux données

Comme mentionné plus haut, le processus MDSV est conçu pour modéliser toute mesure de risque à valeur positive. Il pourrait donc servir à modéliser les variables telles que les variances réalisées, le volume d'actions échangées sur des périodes de 10 minutes ou encore le prix le plus élevé moins le prix le plus bas sur une période de temps spécifique (*high-low range*). Dans le cadre de cette thèse, nous l'utilisons pour modéliser la volatilité des rendements financiers et les variances réalisées d'un actif financier. Nous introduisons, dans cette section, un moyen de relier le processus MDSV, représenté par $V = \{V_t\}, t = 1, \dots, T$, aux rendements financiers et aux variances réalisées.

(1) Rendements financiers

Soit $\{r_t\}, t = 1, \dots, T$, la série temporelle des log-rendements financiers centrés en 0. Comme dans

la plupart des modèles financiers existant dans la littérature, nous modélisons r_t par :

$$r_t = \sqrt{V_t} \epsilon_t, \quad (2.3.1)$$

où $\{\epsilon_t\}$ est un processus d'innovation indépendant et identiquement distribué (*i.i.d.*) de moyenne 0 et de variance 1.

(2) Variances réalisées

Soit $\{RV_t\}$, la série temporelle des variances réalisées qui est obtenue en faisant la somme des carrés des rendements intrajours, telle que présentée à la section 1.1. Puisque cette série est un processus positif, nous la modélisons à l'aide d'une structure d'erreurs multiplicative (MEM) définie par Engle (2002) comme suit :

$$RV_t = V_t \eta_t, \quad (2.3.2)$$

où $\{\eta_t\}$ est un processus d'innovation indépendant et identiquement distribué de moyenne 1. Le modèle MEM est particulièrement performant pour la modélisation des variances réalisées. Par exemple, Engle et Gallo (2006), Lanne (2006), Gallo et Otranto (2015), Augustyniak *et al.* (2019), Cipollini *et al.* (2019a,b) et d'autres auteurs modélisent les variances réalisées avec une structure MEM.

(3) Modèles joints : rendements financiers et variances réalisées

En accord avec l'extension proposée par Augustyniak *et al.* (2019), les modèles factoriels à chaîne de Markov cachée sont très performants pour modéliser et prédire les données financières univariées (rendements financiers ou variances réalisées). Cependant, de récentes études (Takahashi *et al.*, 2009, Hansen *et al.*, 2012, Hansen et Huang, 2016, Liu et Maheu, 2018, Takahashi *et al.*, 2021) montrent qu'au lieu de modéliser uniquement les rendements financiers, le fait d'inclure des mesures réalisées (telles que les variances réalisées ou les variations bi-puissance, *bi-power variations*) permet de mieux capter le niveau actuel de volatilité, une information par la suite utilisée pour former des anticipations sur la volatilité de la prochaine période. Par exemple, les rendements financiers et les variances réalisées peuvent être conjointement modélisés de la manière suivante :

$$r_t = \sqrt{V_t} \epsilon_t, \quad (2.3.3)$$

$$\log RV_t = \xi + \varphi \log V_t + \delta_1 \epsilon_t + \delta_2 (\epsilon_t^2 - 1) + \gamma \varepsilon_t, \quad (2.3.4)$$

où $\xi, \varphi, \delta_1, \delta_2 \in \mathbb{R}$ et $\gamma \in (0, \infty)$ sont des paramètres, et $\{\epsilon_t\}$ et $\{\varepsilon_t\}$ sont des processus d'innovation mutuellement indépendants de moyenne 0 et de variance 1. L'équation (2.3.4) est analogue à celle utilisée par le modèle realized EGARCH de Hansen et Huang (2016) pour modéliser le logarithme des variances réalisées ($\log RV_t$). Dans cette équation, le terme $\delta_1 \epsilon_t + \delta_2 (\epsilon_t^2 - 1)$ crée une dépendance entre $\log RV_t$ et les innovations ϵ_t dirigées par r_t . Comparativement à l'équation (2.3.2), l'équation (2.3.4) apporte plus de flexibilité. À travers le paramètre φ , l'équation (2.3.4) prend en compte le fait que V_t mesure la volatilité au cours de la journée et que RV_t est mesurée durant une partie de la

journée, de l'ouverture à la fermeture de la bourse. L'équation (2.3.4) implique aussi une structure d'erreurs multiplicative (Engle, 2002) sur RV_t . En effet, on peut écrire

$$RV_t = e^\xi \mathbb{E}[\eta_t] V_t^\varphi \frac{\eta_t}{\mathbb{E}[\eta_t]},$$

où $\eta_t = \exp(\delta_1 \epsilon_t + \delta_2 (\epsilon_t^2 - 1) + \gamma \epsilon_t)$ est une variable aléatoire positive.

En se basant sur les résultats de la proposition 2.2.1, on peut déduire les fonctions d'autocovariance et d'autocorrélation du processus d'intérêt correspondant à $\{r_t^2\}$ lorsqu'on modélise uniquement les rendements financiers, à $\{RV_t\}$ lorsqu'on modélise uniquement les variances réalisées et au couple de variables $\{(r_t^2, RV_t)\}$ lorsqu'on modélise conjointement les rendements financiers et les variances réalisées. La proposition 2.3.1 détaille ces fonctions d'autocovariance et d'autocorrélation.

Proposition 2.3.1. (Structure d'autocovariance de $\{r_t^2\}$, $\{RV_t\}$ et $\{(r_t^2, RV_t)\}$) Soit les deux quantités suivantes :

$$\psi_\varphi = \frac{\omega(2 - \nu_0)^{2\varphi} + (1 - \omega)\nu_0^{2\varphi}}{[\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi]^2}, \quad \text{et} \quad \tau_\varphi = \frac{\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi}{[\omega(2 - \nu_0) + (1 - \omega)\nu_0]^\varphi}.$$

Pour $k \in \mathbb{N}$,

(i) si r_t^2 représente le carré des rendements financiers, où r_t est défini par l'équation (2.3.1), alors

$$\begin{aligned} \text{Cov}(r_t^2, r_{t+k}^2) &= \text{Cov}(V_t, V_{t+k}), \\ \text{Var}(r_t^2) &= \sigma^4 \left(\mathbb{E}[\epsilon_t^4] \psi_1^{N(K-1)} - 1 \right). \end{aligned}$$

Si le processus $\{\epsilon_t\}$ est distribué selon une loi normale centrée réduite alors $\mathbb{E}[\epsilon_t^4] = 3$.

(ii) si RV_t représente les variances réalisées, définies par l'équation (2.3.2), alors

$$\begin{aligned} \text{Cov}(RV_t, RV_{t+k}) &= \text{Cov}(V_t, V_{t+k}), \\ \text{Var}(RV_t) &= \sigma^4 \left(\mathbb{E}[\eta_t^2] \psi_1^{N(K-1)} - 1 \right). \end{aligned}$$

Si le processus $\{\eta_t\}$ est distribué selon une loi Gamma ($\text{Gamma}(\nu, 1/\nu)$)⁶ de moyenne 1 et de paramètre de forme ν alors $\mathbb{E}[\eta_t^2] = \frac{\nu + 1}{\nu}$.

(iii) si (r_t, RV_t) est défini par les équations (2.3.3) et (2.3.4), alors

$$\begin{aligned} \text{Cov}(r_t^2, r_{t+k}^2) &= \text{Cov}(V_t, V_{t+k}), \\ \text{Var}(r_t^2) &= \sigma^4 \left(\mathbb{E}[\epsilon_t^4] \psi_1^{N(K-1)} - 1 \right), \end{aligned}$$

et si $\delta_2 < \frac{1}{2}$, on a :

$$\begin{aligned} \text{Cov}(RV_t, RV_{t+k}) &= e^{2\xi} (\mathbb{E}[\eta_t])^2 \text{Cov}(V_t^\varphi, V_{t+k}^\varphi), \\ \text{Var}(RV_t) &= \sigma^{4\varphi} e^{2\xi} \tau_\varphi^{2N(K-1)} \left(\psi_\varphi^{N(K-1)} \mathbb{E}[\eta_t^2] - (\mathbb{E}[\eta_t])^2 \right), \\ \text{Cov}(RV_t, V_t) &= \sigma^{2(\varphi+1)} e^\xi \mathbb{E}[\eta_t] \left(\tau_{\varphi+1}^{N(K-1)} - \tau_\varphi^{N(K-1)} \right). \end{aligned}$$

6. Si $X \sim \text{Gamma}(a, b)$, $a > 0, b > 0$ alors sa moyenne est donnée par $\mathbb{E}[X] = ab$ et sa fonction de densité est donnée par : $g(x | a, b) = (x^{a-1} b^{-a} / \Gamma(a)) \exp(-x/b)$, $x > 0$.

Dans ces équations, $\eta_t = \exp(\delta_1 \epsilon_t + \delta_2 \epsilon_t^2 + \gamma \varepsilon_t - \delta_2)$ et $\text{Cov}(V_t^\varphi, V_{t+k}^\varphi)$ est donné dans la proposition 2.2.1. De plus, si les processus $\{\varepsilon_t\}$ et $\{\epsilon_t\}$ sont normalement distribués alors

$$\mathbb{E}[\eta_t] = \frac{\exp\left(\frac{\delta_1^2}{2-4\delta_2} + \frac{\gamma^2}{2} - \delta_2\right)}{\sqrt{1-2\delta_2}} \quad \text{et} \quad \mathbb{E}[\eta_t^2] = \frac{\exp\left(\frac{2\delta_1^2}{1-4\delta_2} + 2\gamma^2 - 2\delta_2\right)}{\sqrt{1-4\delta_2}}.$$

DÉMONSTRATION. Voir l'annexe A.6. □

2.3.2. Estimation et prévision du modèle MDSV

Désignons par Y_t le processus d'intérêt, c'est-à-dire $Y_t = r_t$ lorsqu'on modélise uniquement les rendements financiers (équation (2.3.1)), $Y_t = RV_t$ lorsqu'on modélise uniquement les variances réalisées (équation (2.3.2)) et $Y_t = (r_t, RV_t)$ lorsqu'on modélise conjointement les rendements financiers et les variances réalisées (équations (2.3.3) et (2.3.4)). Le processus V est une chaîne de Markov homogène ayant un espace d'états de dimension finie (K^N). Ainsi, le processus (Y_t, V_t) est un modèle à chaîne de Markov cachée selon la définition 1.3.5. Le calcul de la log-vraisemblance de ce modèle s'effectue ainsi par les méthodes standards (Hamilton, 1994, chapitre 22). Ainsi, l'algorithme 1 du filtre d'Hamilton présenté à la section 1.3.1 est utilisé pour calculer la vraisemblance dans le modèle MDSV. L'utilisation de cet algorithme nécessite de spécifier les distributions suivantes :

$$\begin{cases} V_1 \sim p_\theta(V_1) \\ V_t | (Y_{1:t-1}, V_{1:t-1}) \sim p_\theta(V_t | V_{t-1}, Y_{1:t-1}), \quad t = 2, 3, \dots \\ Y_t | (Y_{1:t-1}, V_{1:t}) \sim p_\theta(Y_t | V_t, Y_{1:t-1}), \quad t = 1, 2, \dots \end{cases}$$

Dans les applications, nous supposons que V_1 est distribuée selon la distribution stationnaire π de la chaîne de Markov. La matrice de transition de la chaîne de Markov V permet d'obtenir la distribution de V_t conditionnellement à V_{t-1} et aux données passées $Y_{1:t-1}$, notée $p_\theta(V_t | V_{t-1}, Y_{1:t-1})$. La distribution $p_\theta(Y_t | V_t, Y_{1:t-1})$, quant à elle, est donnée par les équations (2.3.1), (2.3.2), (2.3.3), (2.3.4) et la distribution des innovations ϵ_t, η_t et ε_t . Si nous considérons, par exemple, le modèle joint (équations (2.3.3) et (2.3.4)) et que nous supposons que les innovations ϵ_t et ε_t sont distribuées selon une loi normale centrée réduite, alors

$$p_\theta(Y_t | V_t) = p_\theta(r_t, RV_t | V_t) = p_\theta(r_t | V_t) p_\theta(RV_t | V_t, r_t),$$

avec

$$\begin{cases} p_\theta(r_t | V_t) \sim \text{Normal}(0, V_t), \\ p_\theta(RV_t | V_t, r_t) \sim \text{Log-normal}\left(\xi + \varphi \log V_t + \delta_1 \left(\frac{r_t}{\sqrt{V_t}}\right) + \delta_2 \left(\left(\frac{r_t}{\sqrt{V_t}}\right)^2 - 1\right), \gamma^2\right). \end{cases}$$

Il s'agit donc de calculer la log-vraisemblance

$$\mathcal{L} = \log p_\theta(Y_1) + \sum_{t=2}^T \log p_\theta(Y_t | Y_{1:t-1}).$$

Ainsi, l'estimation des paramètres du modèle MDSV s'effectue en maximisant numériquement cette fonction de vraisemblance.

Les techniques de prévision développées par [Hamilton \(1994\)](#) sont aussi applicables dans le cadre du modèle MDSV. Ici, l'intérêt est la prévision des moments de r_{t+h}^2 et RV_{t+h} pour $h = 1, 2, \dots$, conditionnellement à l'information disponible jusqu'à la date t . Les distributions prédictives à l'horizon h de r_{t+h}^2 et RV_{t+h} correspondent à des mélanges de distributions dans lesquels V_{t+h} est la variable de mélange de distribution $\mathbb{P}(V_{t+h} | Y_{1:t})$ où $Y_{1:t} = (Y_1, Y_2, \dots, Y_t)$. Les probabilités conditionnelles de prévision $\mathbb{P}(V_{t+h} | Y_{1:t})$, sont obtenues à partir des probabilités filtrées $\mathbb{P}(V_t | Y_{1:t})$. Les probabilités filtrées, quant à elles, sont obtenues à l'étape 2 du filtre d'Hamilton dont l'algorithme est présenté dans l'algorithme [1](#). Désignons par $\boldsymbol{\xi}_{t|t}$, le vecteur de probabilités filtrées et par $\boldsymbol{\xi}_{t+h|t}$ le vecteur des probabilités conditionnelles de prévision, c'est-à-dire pour tout $h = 0, 1, \dots$,

$$\xi_{j,t+h|t} = \mathbb{P}(V_{t+h} = \nu_j | Y_{1:t}), \quad j = 1, 2, \dots, K^N$$

sont les composantes du vecteur $\boldsymbol{\xi}_{t+h|t}$. Les vecteurs $\boldsymbol{\xi}_{t|t}$ et $\boldsymbol{\xi}_{t+h|t}$ respectent la relation

$$\boldsymbol{\xi}_{t+h|t} = \mathbf{P}^h \boldsymbol{\xi}_{t|t}, \quad h = 1, 2, \dots$$

Par exemple, les prévisions moyennes conditionnelles à l'horizon h de r_{t+h}^2 et RV_{t+h} sont données par :

$$\mathbb{E}[r_{t+h}^2 | Y_{1:t}] = \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j = \boldsymbol{\nu}' \boldsymbol{\xi}_{t+h|t},$$

si le processus $\{r_t\}$ est régi par l'équation [\(2.3.1\)](#) ou par l'équation [\(2.3.3\)](#),

$$\mathbb{E}[RV_{t+h} | Y_{1:t}] = \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j = \boldsymbol{\nu}' \boldsymbol{\xi}_{t+h|t},$$

si le processus $\{RV_t\}$ est régi par l'équation [\(2.3.2\)](#) et

$$\mathbb{E}[RV_{t+h} | Y_{1:t}] = e^\xi \mathbb{E}[\exp(\delta_1 \epsilon_t + \delta_2 (\epsilon_t^2 - 1) + \gamma \epsilon_t)] \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j^\xi.$$

si le processus $\{RV_t\}$ est régi par l'équation [\(2.3.4\)](#). Lorsque les processus $\{\epsilon_t\}$ et $\{\varepsilon_t\}$ dans la dernière équation sont distribués selon une loi normale centrée réduite alors

$$\mathbb{E}[\exp(\delta_1 \epsilon_t + \delta_2 (\epsilon_t^2 - 1) + \gamma \epsilon_t)] = \frac{\exp\left(\frac{\delta_1^2}{2-4\delta_2} + \frac{\gamma^2}{2} - \delta_2\right)}{\sqrt{1-2\delta_2}}.$$

Dans ces équations de prévisions de moyennes conditionnelles à l'horizon h de r_{t+h}^2 et de RV_{t+h} , le symbole $\nu_j, j = 1, 2, \dots, K^N$ désigne le j -ième élément de l'espace d'états $\boldsymbol{\nu}$ de la chaîne de Markov V . Les formules de prévision sont prouvées à l'annexe [A.7](#).

2.4. Étude de simulations

Bien que le nombre d'états dans le modèle MDSV puisse être très grand (K^N), nous pouvons tout de même utiliser le filtre d'Hamilton (section 1.3.1) pour l'estimer. Ce filtre nous permet de calculer la fonction de log-vraisemblance qui a une forme fermée aussi bien pour les modèles MSM, DSARV, CDRS, FHMV que pour le modèle MDSV. La paramétrisation parcimonieuse de la matrice de transition représente la différence principale entre ces modèles et les modèles à chaîne de Markov cachée standards. La fonction de log-vraisemblance calculée par le filtre d'Hamilton est alors maximisée par une routine d'optimisation numérique standard. Notons cependant que l'optimisation numérique de la fonction de vraisemblance du modèle MDSV est de plus en plus lente à mesure que le nombre d'états (K^N) augmente. Pour avoir un temps d'estimation raisonnable, nous nous limitons à des combinaisons de N et K qui donnent un nombre d'états inférieur ou égal à 1024.

Dans cette section, nous présentons quelques propriétés de petits échantillons de l'estimateur du maximum de vraisemblance, évaluées par des simulations de Monte Carlo dans le modèle MDSV. Nous examinons l'estimation des paramètres lorsque le modèle MDSV est spécifié avec un nombre fixe N de composantes et un nombre fixe K d'états des composantes. Pour ce faire, nous simulons et estimons le modèle suivant :

$$r_t = \sqrt{V_t} \epsilon_t,$$

où ϵ_t suit une distribution normale standard et V_t est modélisé par le processus MDSV(N, K) présenté à la section 2.2.

Nous présentons deux panels de simulations : peu de composantes ($N = 2, K = 10$) et beaucoup de composantes ($N = 8, K = 2$). Pour chacun de ces deux panels, les simulations nécessitent cinq paramètres : la variance inconditionnelle σ^2 , la persistance la plus élevée a , le taux de décroissance de la persistance b , le paramètre ν_0 des états et le paramètre de succès ω de la loi stationnaire binomiale. Pour les analyses de simulations, nous fixons les valeurs suivantes sur ces paramètres. La variance inconditionnelle étant juste un facteur d'échelle, nous la normalisons à l'unité $\sigma^2 = 1$; nous utilisons aussi $b = 3$, $a = 0.95$, $\omega \in \{0.2, 0.5, 0.8\}$ et $\nu_0 \in \{0.5, 0.8\}$. La taille de l'échantillon T appartient à $\{T_1, T_2, T_3\}$, où $T_1 = 2500$, $T_2 = 5000$ et $T_3 = 10000$. Le tableau 2.1 rapporte les résultats des simulations Monte Carlo. Les dix-huit (18) colonnes correspondent aux différentes combinaisons des paramètres ω , ν_0 et de la taille de l'échantillon T . Pour chacune des colonnes de chacun des panels, nous avons simulé 1000 échantillons indépendants du modèle approprié (MDSV($N = 2, K = 10$) pour le panel A et MDSV($N = 8, K = 2$) pour le panel B) en utilisant les paramètres et la taille échantillonnale précédemment spécifiés. Nous estimons par la suite les paramètres des modèles par maximum de vraisemblance sur chacune des trajectoires. Pour chaque panel du tableau 2.1, trois (3) lignes concernent chacun des paramètres ($\nu_0, \sigma, a, b, \omega$) : la première ligne donne l'estimation moyenne sur les simulations, la deuxième ligne est l'erreur standard de ces estimations (FSSE, *Finite Sample Standard Error*) et la troisième ligne est la racine carrée de l'écart carré moyen des estimations (RMSE,

TABLE 2.1 – Résultats d'estimation du maximum de vraisemblance des simulations Monte Carlo

| | $\omega = 0.2$ | | | | | | $\omega = 0.5$ | | | | | | $\omega = 0.8$ | | | | | |
|----------------|---------------------------|---------|---------|----------------|---------|---------|----------------|---------|---------|----------------|---------|---------|----------------|---------|---------|----------------|---------|---------|
| | $\nu_0 = 0.50$ | | | $\nu_0 = 0.80$ | | | $\nu_0 = 0.50$ | | | $\nu_0 = 0.80$ | | | $\nu_0 = 0.50$ | | | $\nu_0 = 0.80$ | | |
| | T_1 | T_2 | T_3 | T_1 | T_2 | T_3 | T_1 | T_2 | T_3 | T_1 | T_2 | T_3 | T_1 | T_2 | T_3 | T_1 | T_2 | T_3 |
| | Panel A : $N = 2, K = 10$ | | | | | | | | | | | | | | | | | |
| $\bar{\nu}_0$ | 0,503 | 0,501 | 0,501 | 0,779 | 0,788 | 0,796 | 0,501 | 0,501 | 0,500 | 0,774 | 0,786 | 0,793 | 0,501 | 0,500 | 0,500 | 0,779 | 0,788 | 0,795 |
| FSSE | (0,022) | (0,015) | (0,011) | (0,056) | (0,042) | (0,031) | (0,022) | (0,015) | (0,010) | (0,036) | (0,022) | (0,013) | (0,026) | (0,018) | (0,012) | (0,057) | (0,043) | (0,030) |
| RMSE | (0,022) | (0,015) | (0,011) | (0,060) | (0,044) | (0,031) | (0,022) | (0,015) | (0,010) | (0,045) | (0,026) | (0,015) | (0,026) | (0,018) | (0,012) | (0,060) | (0,044) | (0,030) |
| $\bar{\sigma}$ | 1,032 | 1,019 | 1,009 | 1,011 | 1,002 | 0,999 | 1,048 | 1,009 | 1,011 | 1,002 | 1,001 | 1,002 | 1,013 | 1,005 | 1,007 | 0,998 | 1,000 | 1,002 |
| FSSE | (0,341) | (0,235) | (0,166) | (0,096) | (0,067) | (0,045) | (0,359) | (0,238) | (0,159) | (0,102) | (0,075) | (0,052) | (0,180) | (0,132) | (0,089) | (0,073) | (0,054) | (0,038) |
| RMSE | (0,342) | (0,236) | (0,166) | (0,096) | (0,067) | (0,045) | (0,362) | (0,238) | (0,159) | (0,102) | (0,075) | (0,052) | (0,180) | (0,132) | (0,090) | (0,073) | (0,054) | (0,038) |
| \bar{a} | 0,942 | 0,945 | 0,947 | 0,935 | 0,941 | 0,944 | 0,942 | 0,946 | 0,948 | 0,937 | 0,943 | 0,946 | 0,942 | 0,947 | 0,948 | 0,936 | 0,940 | 0,945 |
| FSSE | (0,021) | (0,016) | (0,010) | (0,026) | (0,021) | (0,016) | (0,021) | (0,015) | (0,009) | (0,025) | (0,018) | (0,013) | (0,021) | (0,014) | (0,010) | (0,027) | (0,020) | (0,015) |
| RMSE | (0,023) | (0,016) | (0,010) | (0,030) | (0,023) | (0,017) | (0,022) | (0,015) | (0,009) | (0,028) | (0,020) | (0,014) | (0,022) | (0,015) | (0,010) | (0,030) | (0,022) | (0,016) |
| \bar{b} | 2,892 | 2,932 | 2,926 | 2,890 | 2,943 | 2,944 | 2,924 | 2,940 | 2,973 | 2,878 | 3,002 | 2,982 | 2,935 | 2,992 | 2,961 | 3,062 | 2,935 | 2,978 |
| FSSE | (1,457) | (1,085) | (0,706) | (2,139) | (1,778) | (1,335) | (1,470) | (1,038) | (0,674) | (1,924) | (1,549) | (1,108) | (1,556) | (1,037) | (0,754) | (2,278) | (1,775) | (1,372) |
| RMSE | (1,461) | (1,087) | (0,710) | (2,142) | (1,779) | (1,336) | (1,472) | (1,039) | (0,675) | (1,928) | (1,549) | (1,108) | (1,557) | (1,037) | (0,755) | (2,279) | (1,776) | (1,372) |
| $\bar{\omega}$ | 0,203 | 0,201 | 0,201 | 0,268 | 0,231 | 0,220 | 0,501 | 0,501 | 0,500 | 0,501 | 0,497 | 0,495 | 0,799 | 0,799 | 0,800 | 0,756 | 0,778 | 0,789 |
| FSSE | (0,020) | (0,012) | (0,009) | (0,206) | (0,142) | (0,096) | (0,026) | (0,017) | (0,010) | (0,218) | (0,171) | (0,129) | (0,026) | (0,015) | (0,009) | (0,182) | (0,126) | (0,083) |
| RMSE | (0,020) | (0,012) | (0,009) | (0,216) | (0,145) | (0,098) | (0,026) | (0,017) | (0,010) | (0,218) | (0,172) | (0,129) | (0,026) | (0,015) | (0,009) | (0,188) | (0,128) | (0,084) |
| | Panel B : $N = 8, K = 2$ | | | | | | | | | | | | | | | | | |
| $\bar{\nu}_0$ | 0,496 | 0,501 | 0,501 | 0,770 | 0,784 | 0,794 | 0,495 | 0,498 | 0,499 | 0,771 | 0,785 | 0,793 | 0,499 | 0,499 | 0,499 | 0,764 | 0,778 | 0,792 |
| FSSE | (0,041) | (0,030) | (0,023) | (0,076) | (0,051) | (0,033) | (0,017) | (0,011) | (0,007) | (0,044) | (0,025) | (0,014) | (0,042) | (0,029) | (0,021) | (0,096) | (0,067) | (0,046) |
| RMSE | (0,042) | (0,030) | (0,023) | (0,081) | (0,053) | (0,034) | (0,018) | (0,011) | (0,008) | (0,053) | (0,029) | (0,016) | (0,042) | (0,029) | (0,021) | (0,103) | (0,070) | (0,047) |
| $\bar{\sigma}$ | 1,005 | 0,999 | 1,000 | 0,999 | 1,001 | 1,001 | 0,994 | 0,996 | 1,000 | 0,998 | 1,000 | 1,001 | 1,002 | 1,001 | 1,000 | 1,002 | 1,000 | 1,000 |
| FSSE | (0,121) | (0,086) | (0,060) | (0,045) | (0,030) | (0,021) | (0,106) | (0,075) | (0,054) | (0,046) | (0,033) | (0,023) | (0,064) | (0,046) | (0,031) | (0,038) | (0,026) | (0,019) |
| RMSE | (0,121) | (0,086) | (0,060) | (0,045) | (0,030) | (0,022) | (0,106) | (0,075) | (0,054) | (0,046) | (0,033) | (0,023) | (0,064) | (0,046) | (0,031) | (0,038) | (0,026) | (0,019) |
| \bar{a} | 0,940 | 0,945 | 0,948 | 0,860 | 0,907 | 0,935 | 0,943 | 0,946 | 0,949 | 0,887 | 0,927 | 0,944 | 0,937 | 0,945 | 0,948 | 0,847 | 0,892 | 0,928 |
| FSSE | (0,031) | (0,019) | (0,013) | (0,176) | (0,108) | (0,047) | (0,023) | (0,015) | (0,010) | (0,135) | (0,068) | (0,026) | (0,033) | (0,018) | (0,011) | (0,188) | (0,130) | (0,073) |
| RMSE | (0,033) | (0,020) | (0,013) | (0,198) | (0,116) | (0,049) | (0,024) | (0,016) | (0,010) | (0,149) | (0,072) | (0,026) | (0,035) | (0,019) | (0,012) | (0,214) | (0,143) | (0,076) |
| \bar{b} | 3,283 | 3,126 | 3,063 | 2,777 | 3,048 | 3,197 | 3,250 | 3,064 | 3,045 | 2,927 | 3,273 | 3,383 | 3,168 | 3,127 | 3,050 | 2,564 | 2,860 | 3,177 |
| FSSE | (1,307) | (0,809) | (0,490) | (1,923) | (1,861) | (1,568) | (1,040) | (0,635) | (0,420) | (1,860) | (1,747) | (1,458) | (1,207) | (0,833) | (0,505) | (1,838) | (1,822) | (1,683) |
| RMSE | (1,337) | (0,819) | (0,494) | (1,936) | (1,862) | (1,580) | (1,070) | (0,638) | (0,423) | (1,862) | (1,769) | (1,508) | (1,219) | (0,843) | (0,508) | (1,889) | (1,827) | (1,693) |
| $\bar{\omega}$ | 0,208 | 0,207 | 0,206 | 0,299 | 0,265 | 0,237 | 0,507 | 0,506 | 0,500 | 0,508 | 0,500 | 0,501 | 0,797 | 0,799 | 0,799 | 0,719 | 0,752 | 0,768 |
| FSSE | (0,064) | (0,046) | (0,032) | (0,240) | (0,192) | (0,127) | (0,080) | (0,057) | (0,041) | (0,227) | (0,177) | (0,127) | (0,043) | (0,028) | (0,020) | (0,242) | (0,185) | (0,128) |
| RMSE | (0,064) | (0,046) | (0,033) | (0,260) | (0,203) | (0,132) | (0,081) | (0,057) | (0,041) | (0,227) | (0,177) | (0,127) | (0,043) | (0,028) | (0,020) | (0,255) | (0,191) | (0,132) |

Chaque panel se base sur $J = 1000$ trajectoires simulées pour chaque colonne. Toutes les simulations du panel A et du panel B se basent sur le processus MDSV avec respectivement ($N = 2, K = 10$) et ($N = 8, K = 2$). Les colonnes se distinguent par les combinaisons de $\omega \in \{0.2, 0.5, 0.8\}$, de $\nu_0 \in \{0.5, 0.8\}$ et de la taille d'échantillon ($T_1 = 2500, T_2 = 5000$, et $T_3 = 10000$). Les paramètres qui sont fixes pour toutes les simulations sont $\sigma = 1, a = 0.95$, et $b = 3$. Sur chaque ensemble de trois lignes, la première ligne est la moyenne des estimations du maximum de vraisemblance sur les J trajectoires simulées. Sur les deux lignes restantes, FSSE est l'erreur standard de l'échantillon et RMSE est l'écart carré moyen des estimations.

Root Mean Square Error). À mesure que la taille de l'échantillon augmente, nous nous attendons à ce que la RMSE et la FSSE se rapprochent.

L'estimation du maximum de vraisemblance donne des résultats raisonnables peu importe le panel. Pour l'estimation des paramètres, le biais est faible et devient négligeable lorsque la taille de l'échantillon augmente. Le paramètre ν_0 a une erreur standard faible relativement à la taille de l'échantillon. Il est donc bien identifié, ce qui est important puisqu'il détermine les fluctuations dans la volatilité. En revanche l'écart-type inconditionnel σ comporte des erreurs standards (FSSE) qui, bien que diminuant comme prévu avec la taille de l'échantillon, représentent environ 10% de la valeur réelle du paramètre. Cela se remarque principalement dans le panel A mais aussi dans le panel B dans une moindre mesure. Ce résultat est cohérent avec les variations de basse fréquence contenues dans les modèles multifractaux tels que le modèle MSM et le modèle MDSV. Cela crée cependant une incertitude considérable sur les moyennes à long terme. Que le modèle ait beaucoup de composantes ($N = 8$) ou peu de composantes ($N = 2$), dans l'ensemble, les simulations Monte Carlo montrent que l'estimation par le maximum de vraisemblance produit des résultats fiables compte tenu des tailles d'échantillon prises en compte.

Nous effectuons une autre analyse par Monte Carlo pour vérifier si la méthode du maximum de vraisemblance permet de sélectionner raisonnablement bien les nombres de composantes N et d'états K . Pour ce faire, nous supposons d'une part que $(N, K) = (5, 2)$ pour tester N (panel C) et d'autre part que $(N, K) = (2, 5)$ pour tester K (panel D). Pour chaque panel, nous simulons 1 000 trajectoires du processus MDSV pour $T = 10\,000$ observations avec les valeurs de N et K correspondantes (MDSV($N = 5, K = 2$) pour le panel C et MDSV($N = 2, K = 5$) pour le panel D) en supposant que $(\nu_0, \sigma, a, b, \omega) = (0.5, 0.8, 0.95, 3, 0.8)$. Pour chaque simulation, nous calculons l'estimateur du maximum de vraisemblance du modèle MDSV(N, K) pour les valeurs de N dans $\{1, 2, \dots, 8\}$ avec $K = 2$ pour le panel C, et pour les valeurs de K dans $\{2, 3, \dots, 9\}$ avec $N = 2$ pour le panel D. Nous sélectionnons ensuite les valeurs de N (pour le panel C) et de K (pour le panel D) qui maximisent la log-vraisemblance. Le tableau 2.2 rapporte le nombre de fois où la valeur N (resp. K) est sélectionnée pour le panel C (resp. panel D). Il rapporte aussi la moyenne et l'erreur standard des différences de log-vraisemblance entre le modèle MDSV($N, 2$) (resp. MDSV($2, K$)) et le vrai modèle MDSV($N = 5, 2$) (resp. MDSV($2, K = 5$)) pour le panel C (resp. panel D).

D'après le panel C dans le tableau 2.2, l'approche du maximum de vraisemblance permet de choisir le modèle avec le vrai nombre de composantes ($N = 5$) ou le modèle avec une composante en moins ($N = 4$) la plupart du temps (501 fois sur 1 000). Nous expliquons cela en remarquant que la composante de volatilité la plus persistante a une fréquence ($1/a$) comparable à la longueur de l'échantillon. Cela fait que sur certaines trajectoires, la composante la plus persistante ne change pas d'état. Le processus MDSV($N = 4, K = 2$) produit alors la log-vraisemblance la plus élevée. Ce résultat reflète le fait général que les événements à basse fréquence peuvent être difficiles à détecter dans un échantillon fini. En conséquence, la méthode d'estimation sous-estime légèrement en moyenne le nombre de composantes. L'analyse du panel D permet de tirer des conclusions similaires pour le nombre d'états K . Le tableau 2.2 indique aussi que le maximum

TABLE 2.2 – Sélection des nombres de composantes et d'états par maximum de vraisemblance

| Panel C : $N = 5, K = 2$ | | | | | | | | |
|------------------------------|---------|---------|---------|---------|---------|---------|---------|---------|
| N | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Fois choisi (max = 1 000) | 0 | 5 | 134 | 284 | 217 | 120 | 70 | 170 |
| $Mean [Lik(N) - Lik(N = 5)]$ | -64.203 | -10.045 | -1.727 | -0.151 | 0.000 | -0.131 | -0.304 | -0.466 |
| $SE [Lik(N) - Lik(N = 5)]$ | (0.430) | (0.160) | (0.067) | (0.023) | - | (0.012) | (0.020) | (0.026) |
| Panel D : $N = 2, K = 5$ | | | | | | | | |
| K | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Fois choisi (max = 1 000) | 0 | 20 | 289 | 283 | 158 | 105 | 57 | 88 |
| $Mean [Lik(K) - Lik(K = 5)]$ | -74.684 | -7.840 | -0.670 | 0.000 | -0.166 | -0.513 | -2.039 | -2.327 |
| $SE [Lik(K) - Lik(K = 5)]$ | (0.603) | (0.168) | (0.045) | - | (0.021) | (0.043) | (0.099) | (0.103) |

Pour chaque simulation du panel C (resp. panel D), le modèle MDSV($N, 2$) (resp. MDSV($2, K$)) est estimé par maximum de vraisemblance pour chacune des valeurs de N entre 1 et 8 (resp. K entre 2 et 9). Pour le panel C (resp. panel D), la première ligne est le nombre de fois que le modèle MDSV($N, 2$) (resp. MDSV($2, K$)) a une log-vraisemblance plus grande que MDSV($5, 2$) (resp. MDSV($2, 5$)), la deuxième ligne donne la moyenne des différences de log-vraisemblances par rapport à la vraie valeur $N = 5$ (resp. $K = 5$) et la troisième ligne donne entre parenthèses les erreurs standards des différences de log-vraisemblances.

de vraisemblance permet de choisir le modèle avec le vrai nombre d'états ($K = 5$) ou le modèle avec un état en moins ($K = 4$) la plupart du temps (572 fois sur 1 000).

Un autre résultat intéressant qui ressort est que lorsque le modèle est sur-paramétrisé (en terme de N et/ou de K), la pénalité est plus faible que lorsqu'il est sous-paramétrisé.

2.5. L'effet levier (*leverage effect*)

Comme précisé au chapitre 1, les séries de rendements financiers présentent parfois un effet levier. Cet effet est particulièrement présent dans les séries de rendements d'indices financiers. Cette caractéristique est très importante pour prédire la volatilité (Corsi et Renò, 2012, Augustyniak *et al.*, 2019, Catania, 2020), pour tarifier des produits dérivés (Han *et al.*, 2020, Wang et Wang, 2021, Takahashi *et al.*, 2021) ou pour gérer le risque financier (Kuester *et al.*, 2006, Chen *et al.*, 2019, Maciel, 2020).

Pour tenir compte de l'effet levier dans le cadre du modèle MDSV, nous utilisons l'approche proposée par Augustyniak *et al.* (2019). Nous modifions donc le processus MDSV en le multipliant par le processus prévisible $\{L_t\}$ comme suit :

$$\begin{aligned}
 V_t &= \sigma^2 \frac{\prod_{i=1}^N C_t^{(i)}}{\prod_{i=1}^N \mathbb{E} [C_t^{(i)}]} L_t, \\
 L_t &= \prod_{i=1}^{N_L} \left\{ 1 + l_i \frac{|r_{t-i}|}{\sqrt{L_{t-i}}} \mathbf{1}_{\{r_{t-i} < 0\}} \right\}, \tag{2.5.1}
 \end{aligned}$$

où $N_L \in \mathbb{N}$ avec $l_1 > 0, l_i = \theta_i^{i-1} l_1$ pour $i = 2, 3, \dots, N_L$ et $\theta_i \in (0, 1)$. Comme chaque composante $C^{(i)}$ du processus MDSV, ce processus d'effet levier $\{L_t\}$ peut être interprété comme étant un ensemble d'effets obtenus à partir des rendements passés. Il autorise qu'un rendement passé ($r_{t-i}, i = 1, 2, \dots, N_L$) négatif ajoute une volatilité supplémentaire d'intensité liée à son amplitude ($|r_{t-i}|$) et à un paramètre l_i structuré de manière à donner de moins en moins d'importance aux rendements les plus éloignés. Par contre, lorsque le rendement passé est positif, il n'a aucun effet sur la volatilité actuelle. Le processus d'effet levier étant prévisible, la valeur de N_L a un impact négligeable sur la charge de calcul et peut donc être choisie comme étant très élevée dans les applications.

L'intégration de l'effet levier au modèle MDSV n'affecte pas sa procédure d'estimation et de prévision à un pas de temps. Cependant, il devient difficile de faire une prévision d'horizon supérieur $h > 1$. En effet, le processus $\{L_t\}$ étant un processus prévisible, c'est-à-dire que L_t est complètement déterminé à partir des données $Y_{1:t-1}$. Ainsi, pour $h = 1$, on a

$$\mathbb{E}[r_{t+1}^2 \mid Y_{1:t}] = L_{t+1} \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j,$$

si le processus $\{r_t\}$ est régi par l'équation (2.3.1) ou par l'équation (2.3.3),

$$\mathbb{E}[RV_{t+1} \mid Y_{1:t}] = L_{t+1} \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j,$$

si le processus $\{RV_t\}$ est régi par l'équation (2.3.2) et

$$\mathbb{E}[RV_{t+1} \mid Y_{1:t}] = e^{\xi} \mathbb{E}[\exp(\delta_1 \epsilon_{t+1} + \delta_2 (\epsilon_{t+1}^2 - 1) + \gamma \epsilon_{t+1})] L_{t+1}^\varphi \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j^\varphi,$$

si le processus $\{RV_t\}$ est régi par l'équation (2.3.4). Par contre, lorsque $h > 1$, il n'existe pas de formule analytique pour calculer les prévisions conditionnelles moyennes. A un horizon supérieur à 1, nous effectuons donc les prévisions par simulation.

Dans le cas où on intègre l'effet levier au modèle MDSV, les fonctions d'autocovariance et d'autocorrélation ne peuvent pas être dérivées analytiquement. Comme pour les prévisions, nous les produisons par simulation.

2.6. Discussion


Les modèles factoriels à chaîne de Markov cachée sont très performants pour modéliser les rendements financiers ou les variances réalisées. Ils supposent que le processus de variance est lié à une chaîne de Markov de grande dimension qui est obtenue par le produit de plusieurs chaînes de Markov à deux états. Dans ce chapitre, nous formalisons le cadre de la modélisation factorielle puis nous introduisons un nouveau processus, le modèle MDSV (*Multifractal Discrete Stochastic Volatility*). Ce dernier apparaît être une généralisation des processus de variance d'autres modèles à chaîne de Markov cachée de haute dimension tels que les modèles

MSM, CDRS, DSARV et FHMV. Notre approche consiste à construire une chaîne de Markov de haute dimension à l'aide d'un produit de chaînes de Markov à deux états ou plus à partir d'une représentation discrète de la volatilité stochastique. Nous considérons aussi un cadre pour modéliser les rendements financiers et les variances réalisées, soit séparément, soit de manière conjointe. L'étude des propriétés et de la structure du processus MDSV permet de remarquer qu'il crée un effet semi-markovien. Cet effet semi-markovien et la flexibilité du modèle MDSV lui permettent de refléter différentes allures de fonctions d'autocorrélation. Le modèle MDSV peut aussi être interprété comme un modèle à volatilité stochastique. Les études de simulations nous montrent que le modèle MDSV peut être raisonnablement bien estimé par le maximum de vraisemblance. De plus, les critères d'information (log-vraisemblance, AIC ou BIC) permettent de bien sélectionner le nombre de composantes N et d'états K du modèle MDSV. Finalement, nous avons observé que la pénalité est plus faible lorsque le modèle est sur-paramétrisé (en terme de N et/ou de K) que lorsqu'il est sous-paramétrisé.

Le chapitre 3 présente une étude empirique du modèle MDSV permettant de mesurer ses performances d'ajustement et de prévision des données financières. Pour ce faire, nous utilisons quatre séries journalières d'indices financiers disponibles sur le site Web de *Oxford-Man Institute Quantitative Finance*, à l'aide desquelles nous effectuons les estimations (avec différentes combinaisons de N et K). Nous étudions aussi les performances du modèle MDSV en terme de gestion du risque.

Chapitre 3

Étude empirique du processus MDSV

Suite au chapitre 2 qui a présenté une analyse théorique du modèle MDSV, le présent chapitre propose une analyse empirique de ce modèle. Il s'agit donc d'étudier à quel point le modèle MDSV s'ajuste aux données financières. D'abord, nous présentons les données financières et les mesures de performance que nous utilisons pour les analyses empiriques. Ensuite, nous effectuons une analyse sous différentes spécifications (différentes combinaisons de N et K) du modèle MDSV afin de sélectionner les valeurs du nombre de composantes N et d'états K utilisées dans la suite du chapitre. Enfin, nous effectuons des comparaisons du modèle MDSV (sous ses formes univariée et jointe) avec des modèles de référence. Pour aller un peu plus loin, nous analysons les avantages comparatifs d'un modèle joint par rapport à un modèle univarié. Nous effectuons aussi une analyse plus approfondie des paramètres estimés par le modèle MDSV joint sur les données d'un indice financier. Pour assurer la reproductibilité des résultats, nous avons créé une librairie (MDSV) sous le logiciel  qui peut être installé à partir du lien web <https://github.com/Abdoulhaki/MDSV> et qui contient tous les codes des résultats obtenus dans ce chapitre.

3.1. Présentation et analyse descriptive des données financières utilisées

L'analyse empirique du modèle MDSV est effectuée sur des séries d'indices financiers disponibles sur le site Web de *Oxford-Man Institute Quantitative Finance*¹. Les données utilisées pour les estimations sont les séries journalières de rendements financiers et de variances réalisées de ces indices sur la période s'étendant du 1^{er} janvier 2000 au 31 décembre 2019. En utilisant les prix à la fermeture de la bourse (*close price*) $\{y_t\}$, nous calculons les log-rendements (*close-to-close log-returns*), exprimés en pourcentage, par la relation

$$r_t = 100(\log y_t - \log y_{t-1}).$$

1. Source des données : <https://realized.oxford-man.ox.ac.uk/data/download>

Cette définition des log-rendements nous permet de capter le risque total sur toute la journée². Rappelons par ailleurs que nous centrons la série des log-rendements et désignons par rendements financiers les log-rendements centrés. Les variances réalisées, quant à elles, sont calculées en utilisant les rendements financiers aux cinq minutes $\{r_{t,i}\}_{i=1,2,\dots,m}$ entre l'ouverture et la fermeture de la bourse, où m désigne le nombre de période de 5 minutes existant entre l'ouverture et la fermeture de la bourse. La variance réalisée dans une journée t est définie par

$$RV_t = 100^2 \sum_{i=1}^m r_{t,i}^2.$$

Les résultats importants que nous présentons dans ce chapitre concernent les quatre indices financiers suivants : S&P 500 (*Standard & Poor's 500*), NASDAQ 100 (*National Association of Securities Dealers Automated Quotation 100*), FTSE 100 (*Financial Times Stock Exchange 100*) et NIKKEI 225 (*Nihon Keizai Shinbun 225*). Les indices S&P 500 et NASDAQ 100 sont les deux indices financiers les plus importants en Amérique du Nord tandis que le FTSE 100 et le NIKKEI 225 sont les indices boursiers les plus importants respectivement en Angleterre et au Japon. En utilisant ces quatre indices, nous avons un aperçu global de la performance du modèle MDSV sur les indices financiers dans le monde. Le tableau 3.1 rapporte les statistiques descriptives des séries journalières de rendements financiers (r_t) et de variances réalisées (RV_t) pour les quatre indices.

TABLE 3.1 – Statistiques descriptives

| Indices | n | Série | Moy. | SD | Asym. | Apla. | Min. | Q1 | Med. | Q3 | Max. |
|------------|------|--------|------|------|-------|--------|--------|-------|------|------|--------|
| S&P 500 | 5016 | r_t | 0.00 | 1.18 | -0.22 | 11.24 | -9.70 | -0.48 | 0.04 | 0.55 | 10.63 |
| | | RV_t | 1.05 | 2.40 | 12.03 | 262.66 | 0.01 | 0.23 | 0.47 | 1.02 | 77.48 |
| NASDAQ 100 | 5012 | r_t | 0.00 | 1.55 | -0.05 | 9.09 | -10.24 | -0.64 | 0.07 | 0.72 | 13.26 |
| | | RV_t | 1.20 | 2.19 | 7.44 | 93.85 | 0.03 | 0.28 | 0.53 | 1.24 | 43.01 |
| FTSE 100 | 5042 | r_t | 0.00 | 1.13 | -0.17 | 9.53 | -8.93 | -0.53 | 0.03 | 0.57 | 9.48 |
| | | RV_t | 1.15 | 2.69 | 17.88 | 543.04 | 0.01 | 0.30 | 0.56 | 1.14 | 106.00 |
| NIKKEI 225 | 4865 | r_t | 0.00 | 1.48 | -0.43 | 9.45 | -12.11 | -0.71 | 0.04 | 0.80 | 13.23 |
| | | RV_t | 1.01 | 1.65 | 8.53 | 111.57 | 0.02 | 0.33 | 0.62 | 1.16 | 32.29 |

Statistiques descriptives des indices S&P 500, NASDAQ 100, FTSE 100, NIKKEI 225 sur la période du 1^{er} janvier 2000 au 31 décembre 2019. n est la taille d'échantillon, Min. est le minimum de la série, Q1 est le premier quartile, Moy. est la moyenne, Med. est la médiane, SD est l'écart-type, Asym. mesure l'asymétrie, Apla. mesure l'aplatissement, Q3 est le troisième quartile et Max. est le maximum.

Les rendements financiers des quatre indices ont une queue épaisse et une asymétrie négative. Ce constat correspond aux faits stylisés (section 1.1.1). Notons que l'asymétrie du NIKKEI 225 est plus prononcée que celle des trois autres indices. En ce qui concerne les variances réalisées, le S&P 500 et le FTSE 100 semblent

2. Il existe d'autres méthodes de calcul des log-rendements dans la littérature. Elles se basent sur le prix à l'ouverture de la bourse (*open-to-close log-returns*), le prix le plus élevé dans la journée de transaction (*open-to-high log-returns*) ou le prix le plus bas dans la journée de transaction (*open-to-low log-returns*).

être plus asymétriques et aplatis que les deux autres indices. Les profils des variances réalisées, présentés à la figure 3.1, montrent une similarité entre ces indices. On remarque aussi des poussées d'activités de marché se produisant en correspondance avec certains événements pertinents : le krach boursier de 2001-2002, la crise financière mondiale de 2008-2009 et le krach boursier de 2015 en Chine. L'effet de ces crises ne semble pas s'estomper rapidement, laissant ainsi présager une dépendance à long terme. En plus, en concordance avec les faits stylisés (section 1.1.2), le regroupement des volatilités se remarque facilement, avec des périodes de faibles volatilités suivies de périodes de fortes volatilités.

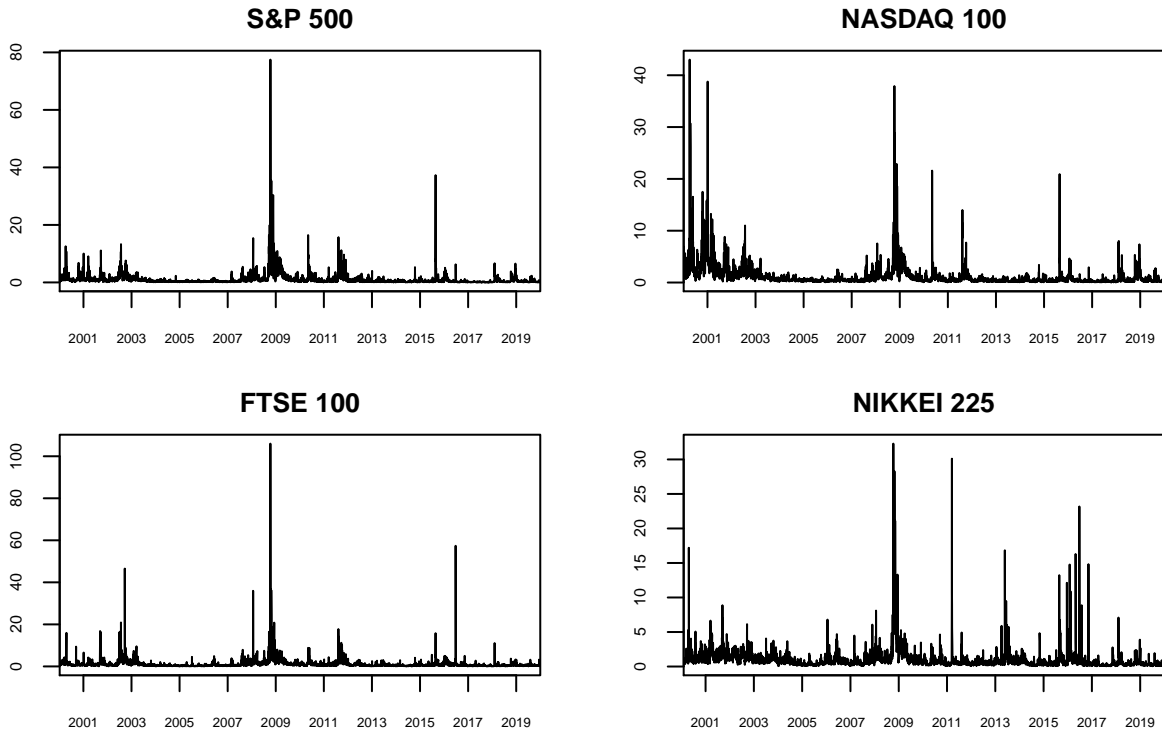


FIGURE 3.1 – Séries de variances réalisées des indices S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225 sur la période du 1^{er} janvier 2000 au 31 décembre 2019.

3.2. Outils de mesures de performance

À travers ce chapitre, nous étudions les performances du modèle MDSV en termes d'ajustement, de prévision et de gestion du risque. Nous utilisons plusieurs outils afin de mesurer ces performances.

3.2.1. Mesures de performance : ajustement et prévision

Il est nécessaire d'avoir des outils pour mesurer les performances d'ajustement et de prévision du modèle MDSV. Ces outils permettent ainsi de comparer les différentes spécifications (différentes combinaisons de N et K) du modèle MDSV entre elles et avec d'autres modèles de référence (*benchmark models*).

3.2.1.1. Mesures de performance : ajustement

Pour mesurer les performances d'ajustement du modèle MDSV nous allons utiliser les critères basés sur la log-vraisemblance maximale atteinte par le modèle. Le calcul de la log-vraisemblance du modèle MDSV s'effectue en utilisant le filtre d'Hamilton présenté dans l'algorithme 1. Outre la log-vraisemblance maximale, que nous désignons par \mathcal{L} , nous utilisons également une transformation des critères d'information d'Akaike (AIC) et bayésien (BIC) standards. Les critères AIC et BIC sont donc calculés à l'aide des formules suivantes :

$$\begin{aligned} \text{AIC} &= \mathcal{L} - k, \\ \text{BIC} &= \mathcal{L} - \frac{k}{2} \log(n), \end{aligned}$$

où k le nombre de paramètres du modèle et n la taille de l'échantillon. Les critères d'information AIC et BIC permettent de pénaliser les modèles en fonction du nombre de paramètres. En effet, puisqu'il est possible d'augmenter la log-vraisemblance d'un modèle en ajoutant un paramètre, cette pénalisation est alors introduite dans le but de satisfaire le critère de parcimonie. Les trois outils utilisés afin de mesurer l'ajustement sont croissants par rapport à la performance. En d'autres termes, plus le critère (log-vraisemblance maximale, AIC ou BIC) est élevé, meilleur est l'ajustement du modèle par rapport aux données.

3.2.1.2. Mesures de performance : prévision

Comme l'estimation, la prévision dans le cadre du modèle MDSV s'effectue par les méthodes standards des modèles à chaîne de Markov cachée (section 2.3.2). Pour l'exercice de prévision, nous prenons les trois dernières années de la base de données, correspondant à 756 jours boursiers, sur lesquelles nous effectuons des prévisions à court et à long terme. Le modèle est ré-estimé à chaque trois mois (63 jours boursiers) pendant la période de prévision. Nous définissons quatre mesures de performance pour la prévision des modèles : la log-vraisemblance prédictive, l'écart quadratique moyen de prévision (RMSFE, *Root Mean Squared Forecast Error*), l'écart absolu moyen de prévision (MAFE, *Mean Absolute Forecast Error*) et la quasi-vraisemblance (QLIK, *Quasi-Likelihood*).

La log-vraisemblance prédictive est la log-densité des rendements financiers futurs étant donné le modèle et l'information disponibles à la date d'évaluation. Dans cette approche, qui a entre autres été considérée par Maheu et McCurdy (2011) et Liu et Maheu (2018), il est question de calculer la log-vraisemblance du rendement du lendemain (r_{t+1}) conditionnellement à l'information disponible actuellement. Le résultat, $\ell_t = \log p(r_{t+1} | \mathcal{F}_t, \theta)$, est utilisé pour définir la log-vraisemblance prédictive sur un horizon de T jours, représentée par

$$d_T = \sum_{t=0}^{T-1} \ell_t = \sum_{t=0}^{T-1} \log p(r_{t+1} | \mathcal{F}_t, \theta).$$

Ici, $t = 0$ correspond à la date finale de l'échantillon d'estimation (*in-sample*), $t = T$ représente la date finale de l'échantillon de prévision (*out-of-sample*) et \mathcal{F}_t désigne toute l'information disponible jusqu'à la date t . Notons qu'une log-vraisemblance prédictive élevée correspond à une meilleure performance de prévision.

L'écart quadratique moyen de prévision (RMSFE), l'écart absolu moyen de prévision (MAFE) et la quasi-vraisemblance (QLIK) nous permettent de mesurer les performances de prévision des modèles aussi bien à court qu'à long terme. Désignons par x_t la variable d'intérêt prédite ($x_t = RV_t$ si la série des variances réalisées est disponible et $x_t = r_t^2$ dans le cas contraire). Les trois critères se basent sur les prévisions cumulatives de la variable d'intérêt $\sum_{i=1}^h \hat{x}_{t+i}$, où $\hat{x}_{t+i} = \mathbb{E}[x_{t+i} | \mathcal{F}_t]$, à différents horizons de temps h . Donc, pour un horizon de temps fixé h , l'écart quadratique moyen de prévision, l'écart absolu moyen de prévision et la quasi-vraisemblance sont respectivement définis par

$$\begin{aligned} \text{RMSFE}(h) &= \sqrt{\frac{1}{T-h+1} \sum_{t=0}^{T-h} \left[\left(\frac{1}{h} \sum_{i=1}^h \hat{x}_{t+i} - \frac{1}{h} \sum_{i=1}^h x_{t+i} \right)^2 \right]}, \\ \text{MAFE}(h) &= \frac{1}{T-h+1} \sum_{t=0}^{T-h} \left| \frac{1}{h} \sum_{i=1}^h \hat{x}_{t+i} - \frac{1}{h} \sum_{i=1}^h x_{t+i} \right|, \\ \text{QLIK}(h) &= \frac{1}{T-h+1} \sum_{t=0}^{T-h} \left[\log \left(\frac{1}{h} \sum_{i=1}^h \hat{x}_{t+i} \right) + \frac{\sum_{i=1}^h x_{t+i}}{\sum_{i=1}^h \hat{x}_{t+i}} \right], \end{aligned}$$

où $t = 0$ correspond à la date finale de l'échantillon d'estimation (*in-sample*) et $t = T$ représente la date finale de l'échantillon de prévision (*out-of-sample*). [Patton \(2011\)](#) a étudié plusieurs fonctions de perte, incluant les fonctions de pertes quadratique, absolue et la quasi-vraisemblance, à partir desquelles sont obtenus les critères RMSFE, MAFE et QLIK. Il a montré que la quasi-vraisemblance est la seule fonction de perte robuste³ qui dépend du rapport (valeur réelle/valeur prédite); également, la fonction de perte quadratique est la seule fonction de perte robuste qui dépend de la différence (valeur réelle – valeur prédite). Les fonctions de pertes quadratique et absolue sont symétriques, c'est-à-dire qu'elles pénalisent une sous-prévision et une sur-prévision de façon équivalente. La quasi-vraisemblance, pour sa part, est asymétrique; cette dernière pénalise davantage une sous-prévision qu'une sur-prévision. La quasi-vraisemblance et les fonctions de pertes quadratique et absolue sont toutes homogènes, de degrés respectifs 0, 1 et 2. Ainsi, si un changement d'unités modifie les données de sorte qu'elles soient multipliées par 100, la quasi-vraisemblance restera inchangée alors que les pertes absolue et quadratique se verront multipliées respectivement par 100 et par 100². Notons aussi que la fonction de perte quadratique est très affectée par les mauvaises prévisions comparativement aux deux autres fonctions de perte.

Ces trois mesures de performance liées à la prévision (l'écart quadratique moyen de prévision, l'écart absolu moyen de prévision et la quasi-vraisemblance) sont décroissantes. En d'autres termes, une faible valeur du critère indique une meilleure performance.

3. Une fonction de perte, est dite « robuste » si les classements de deux prévisions de volatilité (éventuellement imparfaites) par cette fonction de perte sont les mêmes, que le classement soit effectué en utilisant la vraie variance conditionnelle ou une volatilité de substitution (*proxy volatility*) conditionnellement non biaisé.

3.2.2. Mesures de performance : gestion du risque


Au vu des récentes crises financières, particulièrement celle engendrée par la pandémie du COVID-19, la gestion du risque demeure une question primordiale. L'un des outils populaires pour la gestion du risque financier est la valeur à risque (*VaR*, *Value-at-Risk*), qui estime le potentiel de perte selon la probabilité que celle-ci se produise. La *VaR* mesure, à un certain niveau de confiance, le montant maximal de la perte pour un portefeuille d'actifs donné dans des conditions normales de marché sur une période donnée (Jorion, 2000). Mathématiquement, la *VaR* de niveau α , notée $VaR(\alpha)$, est le $100(1 - \alpha)^e$ quantile de la variable $R = -r_t$, c'est-à-dire

$$VaR(\alpha) = -\inf \{x \in \mathbb{R} : \mathbb{P}[r_t \leq x] > \alpha\} = F_R^{-1}(1 - \alpha),$$

où F_R est la fonction de répartition de la variable aléatoire R . La *VaR* peut être interprétée comme une estimation de la queue de la distribution empirique. Dans une approche conditionnelle, la $VaR_{t|t-1}(\alpha)$ est définie sur la distribution conditionnelle des rendements financiers :

$$VaR_{t|t-1}(\alpha) = -\inf \{x \in \mathbb{R} : \mathbb{P}[r_t \leq x | \mathcal{F}_{t-1}] > \alpha\},$$

où \mathcal{F}_{t-1} désigne toute l'information disponible jusqu'à la date $t - 1$.

Dans les modèles à chaîne de Markov cachée, conditionnellement à l'information disponible jusqu'à la date $t - 1$, les rendements financiers r_t sont issus d'un mélange de lois normales. Ainsi, la détermination de la *VaR* dans le cadre du modèle MDSV peut se ramener à un problème de détermination d'un quantile d'un mélange de lois. Cette détermination est facilement accessible en utilisant une méthode numérique (avec la fonction `uniroot` dans le logiciel  par exemple). D'autres méthodes telles que le Monte-Carlo peuvent être utilisées pour déterminer la *VaR* dans un modèle à chaîne de Markov cachée, mais nous optons pour la méthode numérique puisqu'elle est rapide et efficace.

Selon Konderla *et al.* (2017), les événements extrêmes tels que les crises financières sont difficilement capturables par la *VaR*. Celle-ci fournit seulement une évaluation du risque d'investissement dans les conditions normales de marché. Il est alors très important de faire de bonnes prévisions des pertes potentielles. Des procédures de *backtesting* ont été développées pour valider l'utilisation ou non d'un estimateur de la *VaR* (Kupiec, 1995, Christoffersen, 1998, Christoffersen et Pelletier, 2004, Campbell, 2005). Elles consistent principalement à tester deux hypothèses fondamentales portant sur le processus des violations de la *VaR* : l'hypothèse de couverture inconditionnelle et l'hypothèse d'indépendance. Rappelons qu'une violation correspond à une situation dans laquelle le rendement *ex post* est inférieur à la valeur prévue *ex ante* de la *VaR*. En d'autres termes, il y a violation à la date t si le rendement (r_t) observé à la date t est inférieur à la valeur prévue, à la date $t - 1$, de la valeur à risque $VaR_{t|t-1}$. Pour une valeur à risque de niveau α , on s'attend à ce que le taux de violations soit approximativement égal à α . Le test de couverture inconditionnelle, proposé par Kupiec (1995), permet de vérifier si ce taux de violation de la $VaR(\alpha)$ est différent de α . Il stipule tout simplement qu'une *VaR* de niveau α n'est valide qu'à la condition que la fréquence des violations observées ait une espérance égale à $100\alpha\%$. Le test d'indépendance proposé par Christoffersen

(1998), quant à lui, permet de vérifier l'hypothèse d'indépendance des violations. Ce test nous indique s'il y a des violations consécutives de la VaR pendant un certain intervalle de temps. En 2004, [Christoffersen et Pelletier \(2004\)](#) proposent un test conjoint comprenant les deux tests décrits précédemment pour la couverture inconditionnelle et l'indépendance. Ces trois tests sont présentés plus en détails aux sections suivantes.

3.2.2.1. Couverture inconditionnelle : Test de Kupiec

[Kupiec \(1995\)](#) propose un test pour vérifier si la probabilité de violation de la VaR correspond au niveau de confiance spécifié. Soit $I_t(\alpha)$, la variable indicatrice associée à la violation de la $VaR_{t|t-1}(\alpha)$ à la date t , c'est-à-dire

$$I_t(\alpha) = \mathbf{1}_{\{r_t < -VaR_{t|t-1}(\alpha)\}}.$$

L'hypothèse nulle de son test est la suivante :

$$H_0 : \mathbb{P}[I_t(\alpha) = 1] = \mathbb{E}[I_t(\alpha)] = \alpha.$$

La statistique correspondant à ce test est

$$\mathcal{L}_{ci} = 2 \log \left(\frac{\hat{\alpha}^m (1 - \hat{\alpha})^{T-m}}{\alpha^m (1 - \alpha)^{T-m}} \right) \stackrel{as}{\sim} \mathcal{X}_1^2, \quad \text{avec } \hat{\alpha} = \frac{1}{T} \sum_{t=1}^T I_t(\alpha).$$

où $t = 1$ et $t = T$ correspondent respectivement à la date de début et à la date finale de l'échantillon de prévision (*out-of-sample*) et m est le nombre de violations, c'est-à-dire $m = \sum_{t=1}^T I_t(\alpha)$. Notons que le symbole $\stackrel{as}{\sim}$ désigne la convergence asymptotique en distribution. Cette statistique est asymptotiquement distribuée selon une loi du khi carré à un degré de liberté.

3.2.2.2. Indépendance : Test de Christoffersen

Le test proposé par [Christoffersen \(1998\)](#) permet de vérifier l'indépendance des violations de la VaR . Il permet de vérifier que les violations passées de la VaR ne contiennent pas d'informations sur les violations contemporaines et futures. Pour effectuer ce test, on a besoin de déterminer le nombre de violations successives. Désignons respectivement par $n_{00}, n_{01}, n_{10}, n_{11}$ le nombre de non-violations suivies d'une non-violation, le nombre de non-violations suivies d'une violation, le nombre de violations suivies d'une non-violation et le nombre de violations suivies d'une violation. En d'autres termes,

$$\begin{aligned} n_{00} &= \frac{1}{T} \sum_{t=2}^T [1 - I_{t-1}(\alpha)] [1 - I_t(\alpha)], & n_{01} &= \frac{1}{T} \sum_{t=2}^T [1 - I_{t-1}(\alpha)] I_t(\alpha), \\ n_{10} &= \frac{1}{T} \sum_{t=2}^T I_{t-1}(\alpha) [1 - I_t(\alpha)], & \text{et } n_{11} &= \frac{1}{T} \sum_{t=2}^T I_{t-1}(\alpha) I_t(\alpha). \end{aligned}$$

Posons également les probabilités suivantes :

$$\pi_0 = \frac{n_{00}}{n_{00} + n_{01}}, \quad \pi_1 = \frac{n_{11}}{n_{11} + n_{10}}, \quad \pi = \frac{n_{01} + n_{11}}{n_{00} + n_{01} + n_{11} + n_{10}},$$

désignant respectivement la probabilité qu'une violation suive une non-violation, la probabilité qu'une violation suive une violation et la probabilité d'avoir une violation. L'hypothèse nulle du test est la suivante :

$$H_0 : \pi_0 = \pi_1 .$$

La statistique de test satisfait alors

$$\mathcal{L}_{ind} = -2 \log \left(\frac{(1 - \pi)^{n_{00} + n_{10}} \pi^{n_{01} + n_{11}}}{(1 - \pi_0)^{n_{00}} \pi_0^{n_{01}} (1 - \pi_1)^{n_{10}} \pi_1^{n_{11}}} \right) \stackrel{as}{\sim} \mathcal{X}_1^2$$

et suit asymptotiquement une loi du khi carré à un degré de liberté.

3.2.2.3. Couverture conditionnelle

Le test de couverture conditionnelle a été proposé par [Christoffersen et Pelletier \(2004\)](#) et consiste à combiner les tests de couverture inconditionnelle et d'indépendance. La statistique de test

$$\mathcal{L}_{cc} = \mathcal{L}_{ci} + \mathcal{L}_{ind} \stackrel{as}{\sim} \mathcal{X}_2^2$$

est asymptotiquement distribuée selon une loi du khi carré à deux degrés de liberté.

Lorsqu'on rejette les hypothèses nulles des tests de Kupiec et de Christoffersen, le modèle ne permet pas de générer une valeur de la *VaR* qui soit appropriée. De façon équivalente, puisque le test de couverture conditionnelle combine ceux de Kupiec et de Christoffersen, lorsqu'on rejette l'hypothèse nulle de ce test, la *VaR* produite par le modèle ne pourra pas être considérée comme une mesure de risque appropriée.

3.3. Sélection des nombres de composantes et d'états

Le modèle MDSV est lié à deux hyperparamètres (N et K) désignant le nombre de composantes et d'états de la chaîne de Markov. Le choix d'une combinaison de ces deux hyperparamètres définit une spécification de ce modèle. Dans cette section, nous analysons comment les performances du modèle MDSV sont affectées lorsque ces deux hyperparamètres varient. Pour les exercices de sélection de modèle, nous considérons la forme jointe du modèle MDSV (représentée par les équations (2.3.3) et (2.3.4)) qu'on estime à l'aide des données de l'indice financier S&P 500, sans prendre en compte l'effet levier. Dans un premier temps, nous supposons que le modèle MDSV a un nombre fixé de composantes ($N = 1$), puis nous analysons les performances du modèle lorsque le nombre d'états (K) varie de 2 à 1024. Ensuite, nous analysons les performances du modèle lorsque le nombre de composantes (N) varie de 1 à 10 et que le nombre d'états de chaque composante est fixé à $K = 2$. En effet, pour avoir des temps d'estimation raisonnables, nous considérons uniquement des cas où le nombre total d'états (K^N) est inférieur ou égal à 1024.

3.3.1. Sélection du nombre d'états

Les critères de performance montrent que le nombre d'états influence les performances du modèle. En terme d'ajustement, la log-vraisemblance a une tendance croissante à mesure que le nombre d'états augmente,

puis tend à se stabiliser à partir d'un certain K proche de 10 (figure 3.2). La même tendance est constatée lorsqu'on fixe une autre valeur pour le nombre de composantes N ou lorsqu'on effectue l'analyse sur d'autres séries de données financières. Ainsi, pour être efficace, il semblerait que l'utilisation d'un très grand nombre d'états (pour N fixé) ne soit pas nécessaire. En effet, le temps d'estimation du modèle MDSV est lié au nombre total d'états qu'il contient : un plus grand nombre d'états induit un temps d'estimation plus long.

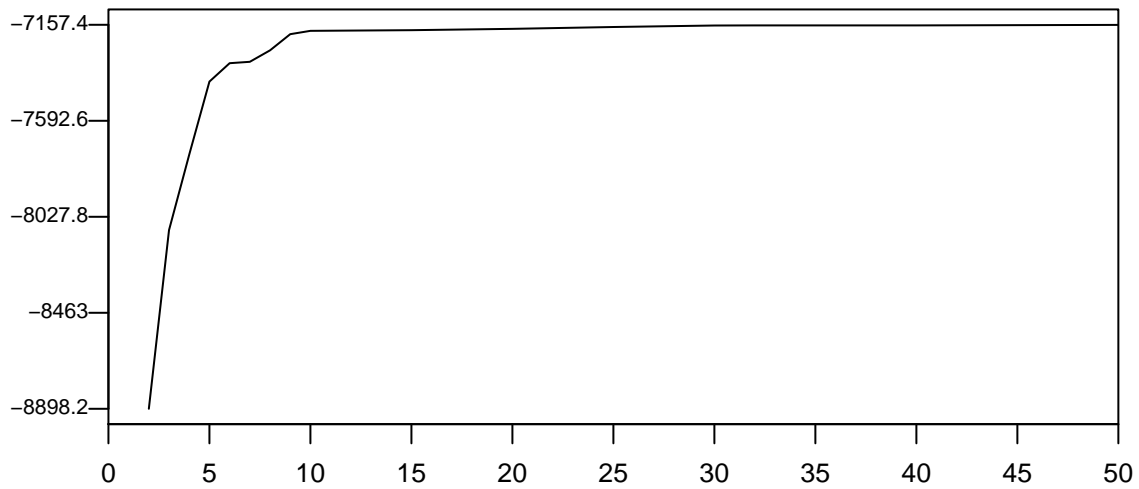


FIGURE 3.2 – Évolution de la log-vraisemblance (en ordonnée) selon le nombre d'états K (en abscisse) dans le modèle MDSV sous sa forme jointe lorsque $N = 1$ sur les données financières S&P 500.

En terme de prévision, on parvient à cette même conclusion. La figure 3.3 permet d'analyser la performance du modèle MDSV sous sa forme jointe lorsque le nombre de composantes est fixé à l'unité et que le nombre d'états augmente. Elle présente les évolutions de la log-vraisemblance prédictive, de l'écart quadratique moyen de prévision à un pas de temps, de la quasi-vraisemblance à un pas de temps et de la valeur-p du test de couverture conditionnelle sur la valeur à risque de niveau 95%. D'après cette figure, la log-vraisemblance prédictive, comme la log-vraisemblance, a une tendance croissante avec le nombre d'états, puis tend à se stabiliser à partir d'un certain K (environ 10). L'écart quadratique moyen de prévision (RM-SFE) et la quasi-vraisemblance (QLIK) à un pas de temps, quant à eux, ont une tendance décroissante par rapport au nombre d'états, puis tendent à se stabiliser à partir d'une certaine valeur de K (environ 10). Ainsi, en terme de prévision, la performance du modèle MDSV sous sa forme jointe s'améliore à mesure qu'on augmente le nombre d'états K , puis tend à se stabiliser à partir d'un certain K . Par souci d'efficacité, il pourrait alors être plus adéquat de ne pas choisir une trop grande valeur pour K .

En terme de gestion du risque, le test de la couverture conditionnelle de niveau 95% ne rejette l'hypothèse nulle en aucun cas puisque la valeur-p est supérieure à 0.05 pour toutes les valeurs de K considérées. De plus à partir de 21 états, le valeur-p est stable aux alentours de 0.4.

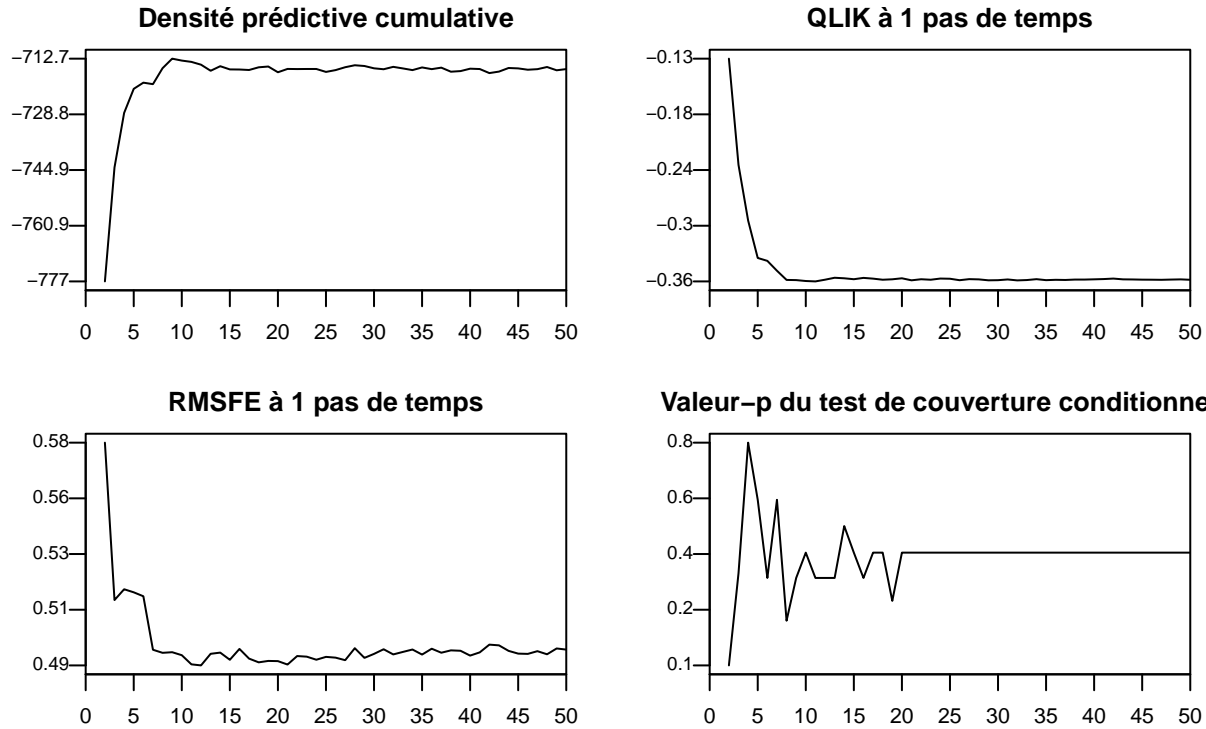


FIGURE 3.3 – Évolution de la log-vraisemblance prédictive, de la quasi-vraisemblance à un pas de temps, de l'écart quadratique moyen de prévision à un pas de temps et de la valeur-p du test de couverture conditionnelle sur la valeur à risque de niveau 95% (en ordonnées) selon le nombre d'états K (en abscisses) dans le modèle MDSV sous sa forme jointe lorsque $N = 1$ sur les données financières de l'indice S&P 500.

3.3.2. Sélection du nombre de composantes

Les performances du modèle MDSV sont aussi affectées par le nombre de composantes N utilisées. Pour un nombre d'états K fixé entre 2 et 6, la log-vraisemblance croît à mesure que le nombre de composantes augmente (figure 3.4). Ainsi, étant donné le nombre d'états, une augmentation du nombre de composantes N permet d'améliorer les performances d'ajustement du modèle MDSV. D'après la figure 3.4, il semble bénéfique de choisir un nombre de composantes N relativement faible en combinaison avec un nombre d'états K relativement élevé. Par exemple, la log-vraisemblance de la spécification ($N = 3, K = 3$) est supérieure à celle de la spécification ($N = 5, K = 2$). De plus, cette dernière comporte 32 états alors que la spécification ($N = 3, K = 3$) en a 27, ce qui constitue un avantage en terme de temps d'estimation.

En terme de prévision, la figure 3.5 présente l'évolution de la log-vraisemblance prédictive, de l'écart quadratique moyen de prévision à un pas de temps et de la quasi-vraisemblance à un pas de temps lorsque

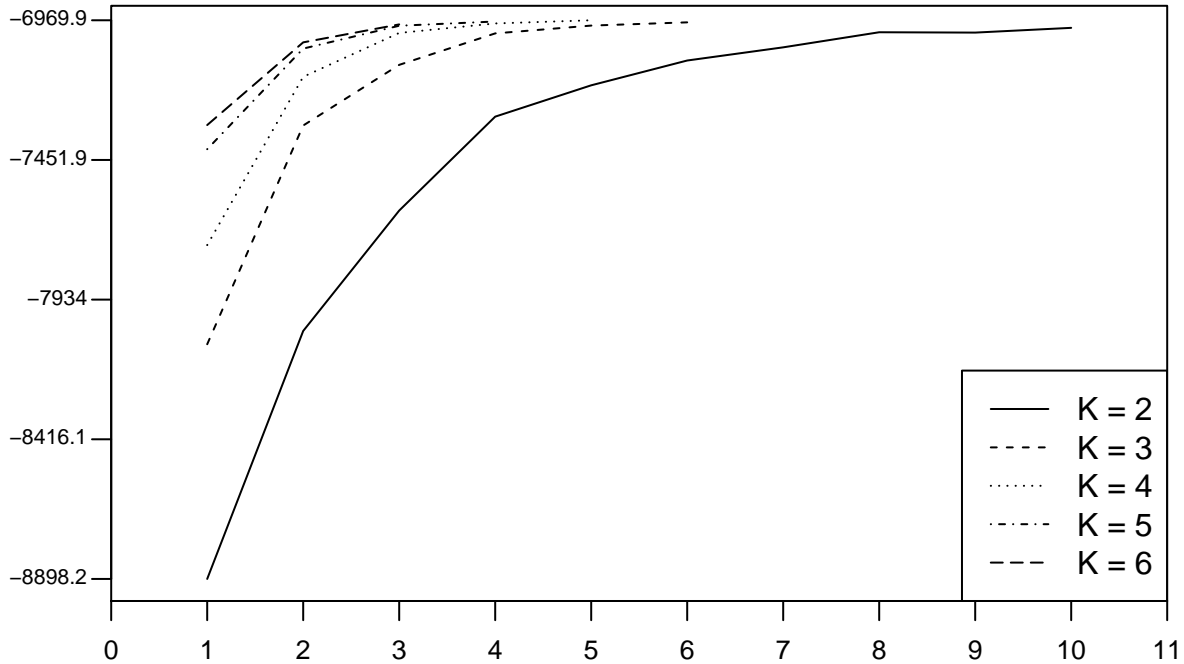


FIGURE 3.4 – Évolution de la log-vraisemblance (en ordonnée) selon le nombre de composantes N (en abscisse) dans le modèle MDSV sous sa forme jointe lorsque $K = 2$ (trait plein), $K = 3$ (tirets), $K = 4$ (pointillés), $K = 5$ (tirets pointillés), $K = 6$ (long tirets), sur les données financières de l'indice S&P 500.

le nombre de composantes augmente. D'après cette figure, la log-vraisemblance prédictive est croissante par rapport au nombre de composantes. Lorsque le nombre de composantes augmente, les critères de performance de prévision RMSFE et QLIK à un pas de temps ont une tendance décroissante. Ainsi, la performance en terme de prévision s'améliore à mesure que le nombre de composantes augmente. La valeur-p du test de couverture conditionnelle sur la valeur à risque de niveau 95% est supérieure à 0.10 pour toutes les valeurs de N , ce qui signifie que même à un niveau de signification égal à 10%, nous ne rejetons jamais l'hypothèse nulle. Ceci nous permet de déduire que la Var à 95% peut être utilisée comme mesure de risque dans le cadre du modèle MDSV.

D'après ces analyses, il est préférable d'utiliser de faibles valeurs pour le nombre d'états ($K < 11$), combinées à un nombre de composantes qui est le plus élevé possible moyennant un temps d'estimation raisonnable. Ainsi, pour les estimations du modèle MDSV et pour limiter le nombre de combinaisons (N, K) , nous considérons les valeurs de N et de K pour lesquelles le nombre total d'états de la chaîne de Markov $\{C_t\}$ est inférieur ou égal à 1024 ($K^N \leq 2^{10} = 1024$). Plus particulièrement, nous présentons dans le cadre de ce chapitre les résultats du modèle MDSV avec les combinaisons de (N, K) suivantes : $(10, 2)$, $(6, 3)$ et $(3, 10)$, correspondant respectivement à un modèle avec beaucoup de composantes, un modèle avec un nombre moyen de composantes et un modèle avec peu de composantes.

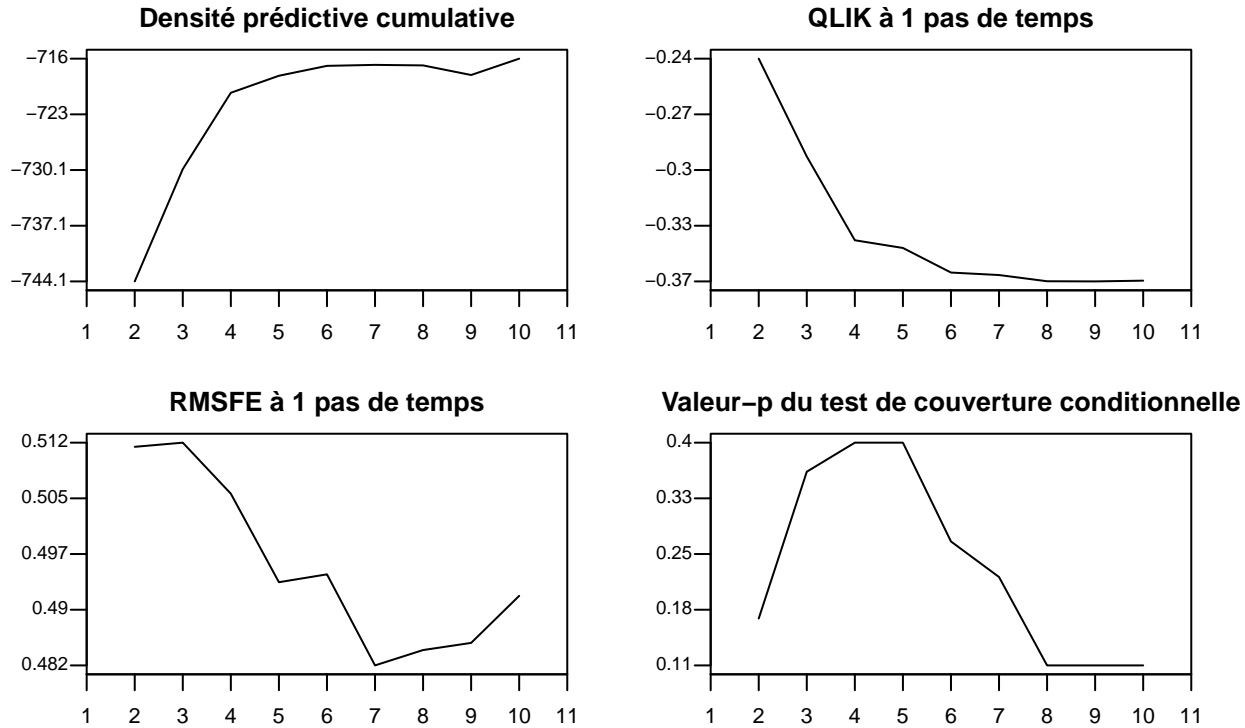


FIGURE 3.5 – Évolution de la log-vraisemblance prédictive, de la quasi-vraisemblance à un pas de temps, de l'écart quadratique moyen de prévision à un pas de temps et de la valeur-p du test de couverture conditionnelle sur la valeur à risque de niveau 95% (en ordonnées) selon le nombre de composantes N (en abscisses) dans le modèle MDSV sous sa forme jointe lorsque chaque composante a $K = 2$ états sur les données financières de l'indice S&P 500.

Dans ce chapitre, nous avons aussi répété les analyses numériques en présence de l'effet levier. À cet effet, nous utilisons le processus effet levier présenté à la section 2.5 avec $N_L = 70$. Nous parvenons à des conclusions similaires pour la sélection du nombre de composantes et du nombre d'états de chaque composante lorsqu'on utilise les résultats des estimations du modèle MDSV tenant compte de l'effet levier.

3.4. Modélisation du processus de rendements financiers

La modélisation univariée des rendements financiers dans le cadre du modèle MDSV s'effectue de façon similaire à d'autres modèles financiers de la littérature. Le rendement financier r_t à la date t vérifie la relation

$$r_t = \sqrt{V_t} \epsilon_t,$$

où $\{\epsilon_t\}$ est un processus d'innovation indépendant et identiquement distribué selon une loi normale de moyenne 0 et de variance 1 et $\{V_t\}$ est le processus MDSV présenté au chapitre 2. Dans cette section, nous analysons empiriquement le modèle MDSV sous sa forme univariée modélisant un processus de rendements financiers. Pour ce faire, nous comparons les performances de trois spécifications du modèle MDSV à celles de modèles de référence. Les trois spécifications utilisées sont celles avec peu de composantes ($N = 3, K = 10$),

avec un nombre moyen de composantes ($N = 6, K = 3$) et avec beaucoup de composantes ($N = 10, K = 2$). Les modèles de référence que nous utilisons sont le modèle GARCH(1,1) tel que présenté à la section 1.2.1, le modèle MSM proposé par [Calvet et Fisher \(2004\)](#) et le modèle FHMV proposé par [Augustyniak et al. \(2019\)](#). Nous utilisons les modèles MSM et FHMV avec 10 composantes; ces modèles sont présentés à la section 1.3.

Dans une deuxième étude comparative tenant compte de l'effet levier, nous remplaçons le modèle GARCH par le modèle GJR de [Glosten et al. \(1993\)](#). Le processus de variance $\{V_t\}$ dans le modèle GJR vérifie la relation

$$V_t = \omega + (\alpha + \gamma \mathbf{1}_{\{r_{t-1} < 0\}}) r_{t-1}^2 + \beta V_t,$$

où $\omega, \alpha, \beta \geq 0$ et $\gamma \in \mathbb{R}$ sont des paramètres du modèle. Notons que le modèle GARCH est un cas particulier du modèle GJR avec $\gamma = 0$. Aussi, étant donné que le modèle MSM n'intègre pas d'effet levier, nous lui en incorporons un grâce au processus d'effet levier multiplicatif présenté à la section 2.5.

3.4.1. Estimation du modèle

Le tableau 3.2 présente les résultats (log-vraisemblance, AIC et BIC) obtenus par la méthode du maximum de vraisemblance pour le modèle MDSV et les modèles de référence (GARCH, MSM et FHMV). Sur chaque ligne du tableau 3.2 (donc pour chaque critère d'efficacité), les valeurs les plus élevées (désignant le meilleur modèle) sont en gras.

En accord avec la littérature sur les données financières, on observe, d'après le tableau 3.2, que les performances de tous les modèles sont largement améliorées lorsqu'on tient compte de l'effet levier. Lorsqu'on ignore l'effet levier, la meilleure spécification MDSV pour modéliser les rendements financiers des indices S&P 500, NASDAQ 100 et NIKKEI 225 est celle avec un nombre moyen de composantes ($N = 6, K = 3$). Pour les données de l'indice FTSE 100, la spécification avec beaucoup de composantes ($N = 10, K = 2$) est préférée. De même, lorsqu'on tient compte de l'effet levier, la meilleure spécification MDSV pour modéliser les rendements financiers des indices NASDAQ 100, FTSE 100 et NIKKEI 225 est celle avec un nombre moyen de composantes ($N = 6, K = 3$). Par contre, pour modéliser les rendements financiers de l'indice S&P 500 en tenant compte de l'effet levier, la spécification avec peu de composantes ($N = 3, K = 10$) est la meilleure. En se basant sur le AIC, le modèle FHMV présente les meilleures performances d'estimation pour tous les indices financiers. Par contre, en se basant sur le BIC, le modèle MDSV est préféré pour modéliser les données de rendements financiers des indices S&P 500, FTSE 100 et le NIKKEI 225. Bien que le modèle MDSV ait moins de paramètres que le modèle FHMV, les performances d'estimation de ces deux modèles sont similaires.

3.4.2. Performance de prévision du modèle

La qualité d'un modèle financier se mesure aussi à sa capacité à bien prévoir la volatilité. Dans cette section, nous analysons les performances de prévision du modèle MDSV sous sa forme univariée modélisant les

TABLE 3.2 – Comparaison des performances d’ajustement : rendements financiers

| Modèles | Sans effet levier | | | | | | Avec effet levier | | | | | |
|-----------------------------------|----------------------|---------|---------|-----------------------|----------------------|-----------------------|----------------------|---------|----------------|-----------------------|----------------------|-----------------------|
| | Modèles de référence | | | MDSV | | | Modèles de référence | | | MDSV | | |
| | GARCH | MSM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | GJR | MSM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| Np | 3 | 4 | 7 | 5 | 5 | 5 | 4 | 6 | 9 | 7 | 7 | 7 |
| S&P 500 ($n = 5016$) | | | | | | | | | | | | |
| \mathcal{L} | -6784.9 | -6678.5 | -6659.8 | -6672.6 | -6668.1 | -6673.1 | -6665.1 | -6593.7 | -6559.4 | -6563.7 | -6606.7 | -6584.5 |
| AIC | -6787.9 | -6682.5 | -6666.8 | -6677.6 | -6673.1 | -6678.1 | -6669.1 | -6599.7 | -6568.4 | -6570.7 | -6613.7 | -6591.5 |
| BIC | -6797.7 | -6695.5 | -6689.7 | -6693.9 | -6689.4 | -6694.4 | -6682.2 | -6619.3 | -6597.7 | -6593.5 | -6636.5 | -6614.3 |
| NASDAQ 100 ($n = 5012$) | | | | | | | | | | | | |
| \mathcal{L} | -8065.6 | -8005.8 | -8000.7 | -8008.3 | -7998.0 | -8011.8 | -7980.6 | -7952.7 | -7934.8 | -7950.0 | -7947.5 | -7947.6 |
| AIC | -8068.6 | -8009.8 | -8007.7 | -8013.3 | -8003.0 | -8016.8 | -7984.6 | -7958.7 | -7943.8 | -7957.0 | -7954.5 | -7954.6 |
| BIC | -8078.3 | -8022.8 | -8030.5 | -8029.6 | -8019.3 | -8033.1 | -7997.6 | -7978.3 | -7973.1 | -7979.9 | -7977.3 | -7977.4 |
| FTSE 100 ($n = 5042$) | | | | | | | | | | | | |
| \mathcal{L} | -6778.5 | -6735.5 | -6738.4 | -6731.5 | -6724.6 | -6722.8 | -6680.3 | -6664.6 | -6654.7 | -6659.0 | -6657.8 | -6657.9 |
| AIC | -6781.5 | -6739.5 | -6745.4 | -6736.5 | -6729.6 | -6727.8 | -6684.3 | -6670.6 | -6663.7 | -6666.0 | -6664.8 | -6664.9 |
| BIC | -6791.3 | -6752.5 | -6768.3 | -6752.8 | -6745.9 | -6744.2 | -6697.3 | -6690.2 | -6693.1 | -6688.9 | -6687.7 | -6687.8 |
| NIKKEI 225 ($n = 4865$) | | | | | | | | | | | | |
| \mathcal{L} | -8195.3 | -8122.7 | -8111.2 | -8111.3 | -8109.0 | -8111.4 | -8145.3 | -8058.0 | -8046.3 | -8061.3 | -8051.0 | -8052.0 |
| AIC | -8198.3 | -8126.7 | -8118.2 | -8116.3 | -8114.0 | -8116.4 | -8149.3 | -8064.0 | -8055.3 | -8068.3 | -8058.0 | -8059.0 |
| BIC | -8208.1 | -8139.7 | -8140.9 | -8132.5 | -8130.2 | -8132.7 | -8162.3 | -8083.5 | -8084.5 | -8091.0 | -8080.7 | -8081.7 |

Np : Nombre de paramètres, \mathcal{L} : log-vraisemblance, AIC : Critère d’information d’Akaike, BIC : Critère d’information bayésien, n : Taille d’échantillon. Les valeurs les plus élevées sont en gras.

rendements financiers. Le tableau 3.3 présente la log-vraisemblance prédictive d_T pour les trois spécifications du modèle MDSV et pour les modèles de référence (GARCH/GJR GARCH, MSM et FHMV). Pour chaque indice, les valeurs les plus élevées (désignant le meilleur modèle) sont en gras.

TABLE 3.3 – Log-vraisemblance prédictive : rendements financiers

| Modèles | Sans effet levier | | | | | | Avec effet levier | | | | | |
|------------|----------------------|--------|--------|-----------------------|----------------------|-----------------------|----------------------|--------|---------------|-----------------------|----------------------|-----------------------|
| | Modèles de référence | | | MDSV | | | Modèles de référence | | | MDSV | | |
| | GARCH | MSM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | GJR | MSM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| S&P 500 | -787.7 | -744.7 | -739.8 | -742.7 | -742.3 | -744.0 | -766.3 | -728.3 | -722.8 | -735.3 | -734.5 | -729.1 |
| NASDAQ 100 | -985.3 | -971.0 | -958.3 | -958.1 | -957.9 | -969.2 | -971.7 | -952.6 | -942.9 | -951.3 | -950.4 | -950.6 |
| FTSE 100 | -813.9 | -798.5 | -801.0 | -798.3 | -797.3 | -797.1 | -815.0 | -796.9 | -791.7 | -796.2 | -796.6 | -795.6 |
| NIKKEI 225 | -1023.1 | -993.4 | -990.4 | -991.6 | -990.2 | -990.6 | -1011.1 | -973.5 | -974.8 | -983.9 | -976.5 | -972.8 |

Les valeurs les plus élevées sont en gras.

La log-vraisemblance prédictive mesure à quel point le rendement du lendemain est probable compte tenu de l'information disponible jusqu'à présent. D'après le tableau 3.3, prendre en compte l'effet levier permet d'améliorer la qualité des prévisions puisque la log-vraisemblance prédictive de chaque modèle augmente lorsque celui-ci intègre l'effet levier. Qu'on considère le cas où l'effet levier est pris en compte ou le cas contraire, les performances des modèles sont similaires dans chacun de ces deux sous-groupes. Dans quatre cas sur huit, le modèle MDSV présente les meilleures performances, alors que le modèle FHMV présente les meilleures performances dans les quatre cas restants. En effet, lorsqu'on ne tient pas compte de l'effet levier, le modèle MDSV présente les meilleures densités prédictives sur les données des indices financiers NASDAQ 100, FTSE 100 et NIKKEI 225 alors que le modèle FHMV est préféré sur les données du S&P 500. Par contre, lorsqu'on tient compte de l'effet levier, le modèle FHMV présente les densités prédictives les plus élevées pour les indices financiers S&P 500, NASDAQ 100 et FTSE 100. Le modèle MDSV, quant à lui, est le meilleur en terme de log-vraisemblance prédictive pour l'indice NIKKEI 225. Parmi les spécifications du modèle MDSV, la spécification avec beaucoup de composantes ($N = 10, K = 2$) est la meilleure en terme de log-vraisemblance prédictive pour les indices financiers S&P 500, FTSE 100 et NIKKEI 225 alors que la spécification avec un nombre moyen de composantes ($N = 6, K = 3$) est meilleure pour le NASDAQ 100.

Les écarts quadratique et absolu moyens, ainsi que la quasi-vraisemblance, nous permettent de mesurer la performance de prévision des modèles à différents horizons de temps. Le tableau 3.4 présente les $RMSFE(h)$, $MAFE(h)$ et $QLIK(h)$ des modèles de référence et du modèle MDSV pour les horizons $h = 1, 5, 25$ et 100 correspondant respectivement à une prévision d'un jour, d'une semaine, d'environ un mois et d'environ quatre mois. Pour chaque indice, chaque mesure de performance et chaque horizon (h), les valeurs les plus faibles (désignant le meilleur modèle) sont en gras. La signification statistique des différences entre les mesures de performance est évaluée par le test de Diebold-Mariano (Diebold et Mariano, 2002). Pour ce test, nous utilisons le modèle FHMV comme modèle de référence.

Le tableau 3.4 nous permet de constater que l'effet levier permet d'améliorer les prévisions à court terme (horizon d'une semaine ou moins) pour chacun des modèles. Par contre, pour les horizons plus longs, tenir compte de l'effet levier n'a généralement pas d'effet et peut même, dans certains cas, être nuisible. D'après le tableau 3.4, dans 91 cas sur 144 (horizon de plus d'une semaine), l'effet levier a intérêt insignifiant ou est nuisible. En effet, les erreurs commises dans les prévisions s'accumulent et puisque l'effet levier est lié aux rendements passés, le fait d'en tenir compte conduit à une plus grande erreur de prévision que le fait de l'ignorer. Lorsqu'on ne tient pas compte de l'effet levier, le modèle MDSV présente des critères significativement (à 1%, 5% ou 10%) plus faibles que le modèle FHMV à court terme (23 cas sur 24). Par contre, en tenant compte de l'effet levier, le modèle MDSV ne présente pas un gain significatif par rapport au FHMV à court terme. De façon générale, à long terme, le modèle MDSV présente les critères les plus faibles dans 19 cas sur 24. Plus particulièrement, en se basant sur le QLIK, le modèle MDSV produit les meilleures prévisions à long terme sur les indices S&P 500, FTSE 100 et NIKKEI 225 tandis que le modèle GARCH produit les meilleures prévisions à long terme sur le NASDAQ 100. À court terme, le MDSV fait

TABLE 3.4 – Performance de prévision de volatilité : rendements financiers

| | | RMSFE | | | | MAFE | | | | QLIK | | | |
|-------------------|-------------|--------------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| | | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 |
| | | S&P 500 | | | | | | | | | | | |
| Sans effet levier | GARCH | 1.65* | 0.95 | 0.75 | 0.67 | 0.75 | 0.53 | 0.54 | 0.54 | 0.25 | 1.94 | 3.73 | 5.33 |
| | MSM | 1.65‡ | 0.93† | 0.70† | 0.60‡ | 0.72‡ | 0.50‡ | 0.49† | 0.47‡ | 0.25 | 1.95 | 3.76 | 5.37 |
| | FHVM | 1.66 | 0.96 | 0.77 | 0.76 | 0.74 | 0.52 | 0.52 | 0.60 | 0.26 | 1.96 | 3.77 | 5.36 |
| | MDSV(3, 10) | 1.65 | 0.94 | 0.73 | 0.65‡ | 0.73* | 0.50* | 0.51 | 0.52‡ | 0.25† | 1.94 | 3.76 | 5.37 |
| | MDSV(6, 3) | 1.65* | 0.94† | 0.71* | 0.60‡ | 0.71‡ | 0.48‡ | 0.48† | 0.48‡ | 0.24‡ | 1.94‡ | 3.78 | 5.45 |
| | MDSV(10, 2) | 1.65† | 0.93* | 0.69* | 0.58‡ | 0.70‡ | 0.47‡ | 0.46† | 0.44‡ | 0.24‡ | 1.94* | 3.80 | 5.51 |
| Avec effet levier | GJR | 1.61 | 0.94 | 0.83 | 0.92 | 0.75 | 0.53 | 0.60 | 0.75 | 0.19 | 1.91 | 3.76 | 5.40 |
| | MSM | 1.59 | 0.88 | 0.67 | 0.53† | 0.70 | 0.46 | 0.48 | 0.43* | 0.18 | 1.88 | 3.76 | 5.34 |
| | FHVM | 1.58 | 0.88 | 0.72 | 0.86 | 0.69 | 0.47 | 0.49 | 0.65 | 0.17 | 1.88 | 3.76 | 5.47 |
| | MDSV(3, 10) | 1.62 | 1.03 | 1.05 | 1.36 | 0.75 | 0.54 | 0.63 | 0.93 | 0.16 | 1.87 | 3.76 | 5.44 |
| | MDSV(6, 3) | 1.59 | 0.92 | 0.74 | 0.72† | 0.70 | 0.48 | 0.49 | 0.55 | 0.16 | 1.88 | 3.75 | 5.40 |
| | MDSV(10, 2) | 1.61 | 0.89 | 0.68 | 0.54† | 0.70 | 0.46 | 0.49 | 0.45* | 0.15 | 1.86 | 3.72 | 5.31 |
| | | NASDAQ 100 | | | | | | | | | | | |
| Sans effet levier | GARCH | 2.10† | 1.14† | 0.91† | 0.87‡ | 1.10‡ | 0.73‡ | 0.68* | 0.72‡ | 0.77† | 2.42* | 4.17 | 5.74† |
| | MSM | 2.15 | 1.24 | 1.04 | 1.06‡ | 1.17 | 0.79 | 0.73 | 0.78‡ | 0.82 | 2.46 | 4.20 | 5.81 |
| | FHVM | 2.11 | 1.19 | 1.05 | 1.45 | 1.13 | 0.77 | 0.75 | 1.10 | 0.79 | 2.45 | 4.20 | 5.81 |
| | MDSV(3, 10) | 2.09† | 1.14‡ | 0.90† | 0.84‡ | 1.09‡ | 0.72‡ | 0.66† | 0.67‡ | 0.78 | 2.44 | 4.18 | 5.75‡ |
| | MDSV(6, 3) | 2.09‡ | 1.12‡ | 0.88† | 0.77‡ | 1.07‡ | 0.68‡ | 0.61‡ | 0.57‡ | 0.77† | 2.43* | 4.19 | 5.77 |
| | MDSV(10, 2) | 2.09† | 1.13† | 0.89† | 0.76‡ | 1.09‡ | 0.70‡ | 0.66* | 0.60‡ | 0.77 | 2.44 | 4.22 | 5.77‡ |
| Avec effet levier | GJR | 2.08 | 1.15 | 1.00 | 1.10 | 1.11 | 0.75 | 0.77 | 0.90 | 0.73 | 2.42 | 4.19 | 5.78 |
| | MSM | 2.06 | 1.12 | 0.91 | 0.95 | 1.05 | 0.68 | 0.66 | 0.75 | 0.72 | 2.40 | 4.18 | 5.77 |
| | FHVM | 2.07 | 1.12 | 0.92 | 0.94 | 1.03 | 0.67 | 0.66 | 0.71 | 0.72 | 2.40 | 4.19 | 5.81 |
| | MDSV(3, 10) | 2.08 | 1.15 | 1.00 | 1.23 | 1.09 | 0.73 | 0.72 | 0.95 | 0.73 | 2.41 | 4.19 | 5.80 |
| | MDSV(6, 3) | 2.06 | 1.11 | 0.88* | 0.83* | 1.04 | 0.67 | 0.62* | 0.64 | 0.72 | 2.39 | 4.18 | 5.77 |
| | MDSV(10, 2) | 2.06 | 1.11* | 0.87* | 0.77† | 1.02 | 0.65† | 0.60† | 0.57† | 0.71* | 2.39* | 4.18 | 5.78 |
| | | FTSE 100 | | | | | | | | | | | |
| Sans effet levier | GARCH | 1.07‡ | 0.60‡ | 0.41† | 0.44† | 0.60‡ | 0.41† | 0.33‡ | 0.38† | 0.32‡ | 1.96† | 3.61 | 5.10† |
| | MSM | 1.06‡ | 0.58‡ | 0.37‡ | 0.33‡ | 0.58‡ | 0.39‡ | 0.30‡ | 0.26‡ | 0.32‡ | 1.96‡ | 3.61‡ | 5.04‡ |
| | FHVM | 1.08 | 0.61 | 0.44 | 0.54 | 0.61 | 0.42 | 0.35 | 0.45 | 0.36 | 1.99 | 3.63 | 5.13 |
| | MDSV(3, 10) | 1.07‡ | 0.59‡ | 0.38‡ | 0.32‡ | 0.59‡ | 0.40‡ | 0.30‡ | 0.26‡ | 0.32‡ | 1.96‡ | 3.60‡ | 5.04‡ |
| | MDSV(6, 3) | 1.07‡ | 0.59‡ | 0.37‡ | 0.31‡ | 0.58‡ | 0.39‡ | 0.29‡ | 0.24‡ | 0.31‡ | 1.95‡ | 3.60‡ | 5.03‡ |
| | MDSV(10, 2) | 1.06‡ | 0.59‡ | 0.36‡ | 0.30‡ | 0.58‡ | 0.39‡ | 0.29‡ | 0.22‡ | 0.31‡ | 1.95‡ | 3.60‡ | 5.02‡ |
| Avec effet levier | GJR | 1.08 | 0.64 | 0.50 | 0.57* | 0.61 | 0.43 | 0.38 | 0.50* | 0.32 | 1.98 | 3.64 | 5.16 |
| | MSM | 1.05 | 0.59 | 0.39 | 0.36* | 0.57 | 0.39 | 0.31* | 0.27† | 0.29 | 1.95 | 3.61* | 5.05‡ |
| | FHVM | 1.06 | 0.62 | 0.81 | 2.92 | 0.57 | 0.39 | 0.37 | 1.14 | 0.29 | 1.95 | 3.63 | 5.28 |
| | MDSV(3, 10) | 1.06 | 0.62 | 0.43 | 0.35* | 0.58 | 0.40 | 0.31* | 0.27† | 0.29 | 1.95 | 3.61* | 5.05† |
| | MDSV(6, 3) | 1.06 | 0.60 | 0.40 | 0.35* | 0.58 | 0.39 | 0.30* | 0.26† | 0.30 | 1.95 | 3.60† | 5.04‡ |
| | MDSV(10, 2) | 1.05 | 0.58 | 0.37 | 0.34* | 0.57 | 0.38 | 0.28† | 0.25† | 0.29 | 1.94 | 3.59† | 5.04‡ |
| | | NIKKEI 225 | | | | | | | | | | | |
| Sans effet levier | GARCH | 2.14 | 1.23 | 0.95 | 1.19 | 1.17 | 0.84 | 0.76 | 1.05 | 0.87 | 2.51 | 4.20 | 5.78 |
| | MSM | 2.13‡ | 1.18* | 0.83† | 0.91‡ | 1.11‡ | 0.77‡ | 0.62‡ | 0.77‡ | 0.86 | 2.49 | 4.15 | 5.69‡ |
| | FHVM | 2.14 | 1.20 | 0.88 | 1.04 | 1.14 | 0.80 | 0.65 | 0.89 | 0.86 | 2.49 | 4.16 | 5.72 |
| | MDSV(3, 10) | 2.13 | 1.19 | 0.83† | 0.86‡ | 1.10‡ | 0.77† | 0.60‡ | 0.72‡ | 0.85 | 2.49 | 4.15 | 5.68‡ |
| | MDSV(6, 3) | 2.12† | 1.16* | 0.76‡ | 0.67‡ | 1.06‡ | 0.73‡ | 0.52‡ | 0.57‡ | 0.85 | 2.49 | 4.14 | 5.64† |
| | MDSV(10, 2) | 2.12‡ | 1.16† | 0.75‡ | 0.66‡ | 1.07‡ | 0.73‡ | 0.53‡ | 0.57‡ | 0.85 | 2.49 | 4.14 | 5.63† |
| Avec effet levier | GJR | 2.14 | 1.25 | 1.07 | 1.38 | 1.17 | 0.85 | 0.86 | 1.24 | 0.84 | 2.50 | 4.23 | 5.83 |
| | MSM | 2.12 | 1.11 | 0.74* | 0.72‡ | 1.07* | 0.72† | 0.58 | 0.62‡ | 0.80 | 2.46 | 4.15 | 5.65‡ |
| | FHVM | 2.08 | 1.11 | 0.79 | 0.90 | 1.09 | 0.74 | 0.61 | 0.76 | 0.80 | 2.45 | 4.15 | 5.69 |
| | MDSV(3, 10) | 2.10 | 1.15 | 0.80 | 0.84† | 1.06‡ | 0.74 | 0.58 | 0.71† | 0.81 | 2.47 | 4.15 | 5.68 |
| | MDSV(6, 3) | 2.12 | 1.10 | 0.73 | 0.64‡ | 1.05‡ | 0.69‡ | 0.53† | 0.57† | 0.80 | 2.47 | 4.16 | 5.64 |
| | MDSV(10, 2) | 2.11 | 1.10 | 0.72 | 0.63‡ | 1.03‡ | 0.67‡ | 0.49‡ | 0.55† | 0.80 | 2.46 | 4.15 | 5.66 |

Les symboles *, † et ‡ signifient respectivement que le critère est significativement inférieur à celui du modèle de référence (FHVM) aux seuils de 10%, 5% et 1% en utilisant le test de Diebold-Mariano.

de meilleures prévisions sur les indices financiers S&P 500 et NASDAQ 100 tandis que le modèle FHMV produit les meilleures prévisions à court terme sur le NIKKEI 225. Si on se base sur le MAFE, on conclut que le MDSV produit les meilleures prévisions aussi bien à court qu'à long terme pour le NASDAQ 100 et le NIKKEI 100. La même conclusion est tirée sur les données de rendements financiers du NASDAQ 100 et du FTSE 100 lorsqu'on se base sur le RMSFE. Notons que les performances du modèle s'améliorent à mesure que l'horizon de prévision croît. Par exemple les RMSFE et MAFE du modèle MDSV à l'horizon $h = 100$ jours sont significativement inférieurs à ceux du FHMV pour tous les indices financiers.

Parmi les trois spécifications du modèle MDSV, celle avec beaucoup de composantes ($N = 10, K = 2$) semble mener au modèle qui produit les meilleures prévisions, aussi bien à court qu'à long terme. Cela laisse penser que plus le nombre de composantes est élevé, meilleures sont les prévisions. Une autre remarque importante est que le modèle MDSV avec beaucoup de composantes ($N = 10, K = 2$) semble produire les meilleures prévisions sur un horizon d'un mois pour tous les indices (en se basant sur le MAFE).

3.4.3. Performance de gestion du risque du modèle

Pour mesurer les performances de gestion du risque sous le modèle MDSV univarié modélisant les rendements financiers, nous utilisons les tests de couverture inconditionnelle, d'indépendance et de couverture conditionnelle basés sur la valeur à risque de niveau 95%. Le tableau 3.5 présente les valeurs-p de ces trois tests, pour chacun des modèles de référence et des trois spécifications du modèle MDSV. Les valeurs-p inférieures aux seuils de 5% sont marquées par le symbole *.

D'après le tableau 3.5, à part le modèle GARCH qui donne, dans certains cas, des statistiques permettant de rejeter l'hypothèse nulle, tous les modèles MDSV produisent une bonne mesure de la valeur à risque à 95%. De plus, le fait de tenir compte de l'effet levier dans les spécifications avec un nombre moyen de composantes ($N = 6, K = 3$) et avec beaucoup de composantes ($N = 10, K = 2$) du modèle MDSV permet de produire des statistiques de tests plus faibles. Ainsi, en utilisant le modèle MDSV ou les modèles FHMV et MSM pour la gestion du risque financier, on prévoit que les pertes quotidiennes dépasseront la valeur à risque prédite dans environ 5% des cas. Parmi les trois spécifications du modèle MDSV, le modèle avec beaucoup de composantes ($N = 10, K = 2$) produit la valeur-p la plus élevée (statistique la plus faible) pour le test de couverture conditionnelle sur le S&P 500 alors que celle avec un nombre moyen de composantes ($N = 6, K = 3$) produit les valeurs-p les plus élevées (statistiques les plus faibles) pour le test de couverture conditionnelle sur les indices financiers NASDAQ 100 et NIKKEI 225.

3.5. Modélisation du processus de variances réalisées

En finance, une autre mesure est souvent utilisée pour estimer la variance du prix d'un actif sur une période de temps ; il s'agit de la variance réalisée. La variance réalisée fournit une mesure assez précise de la volatilité d'un actif. Elle est donc souvent utilisée pour faire de la prévision de volatilité ou pour mesurer les performances de prévision. Dans la littérature, les modèles financiers les plus connus pour la modélisation

TABLE 3.5 – Performance de gestion du risque : rendements financiers

| Modèles | Sans effet levier | | | | | | Avec effet levier | | | | | |
|---------------------|----------------------|-------|-------|-------------------|------------------|-------------------|----------------------|-------|-------|-------------------|------------------|-------------------|
| | Modèles de référence | | | MDSV | | | Modèles de référence | | | MDSV | | |
| | GARCH | MSM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | GJR | MSM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| S&P 500 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.037* | 0.842 | 0.762 | 0.636 | 0.598 | 0.395 | 0.037* | 0.842 | 0.519 | 0.321 | 0.636 | 0.842 |
| \mathcal{L}_{ind} | 0.010* | 0.187 | 0.112 | 0.304 | 0.599 | 0.329 | 0.292 | 0.993 | 0.265 | 0.586 | 0.762 | 0.993 |
| \mathcal{L}_{cc} | 0.004* | 0.410 | 0.270 | 0.528 | 0.758 | 0.433 | 0.066 | 0.980 | 0.437 | 0.527 | 0.854 | 0.980 |
| NASDAQ 100 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.242 | 0.178 | 0.519 | 0.893 | 0.395 | 0.716 | 0.087 | 0.842 | 0.762 | 0.519 | 0.893 | 0.842 |
| \mathcal{L}_{ind} | 0.531 | 0.115 | 0.702 | 0.884 | 0.329 | 0.544 | 0.970 | 0.993 | 0.823 | 0.634 | 0.884 | 0.993 |
| \mathcal{L}_{cc} | 0.415 | 0.117 | 0.755 | 0.981 | 0.433 | 0.778 | 0.231 | 0.980 | 0.932 | 0.725 | 0.981 | 0.980 |
| FTSE 100 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.126 | 0.762 | 0.636 | 0.893 | 0.716 | 0.973 | 0.087 | 0.893 | 0.893 | 0.716 | 0.716 | 0.973 |
| \mathcal{L}_{ind} | 0.118 | 0.112 | 0.093 | 0.134 | 0.070 | 0.045* | 0.380 | 0.392 | 0.392 | 0.070 | 0.070 | 0.159 |
| \mathcal{L}_{cc} | 0.092 | 0.270 | 0.218 | 0.322 | 0.180 | 0.135 | 0.157 | 0.687 | 0.687 | 0.180 | 0.180 | 0.370 |
| NIKKEI 225 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.037* | 0.636 | 0.413 | 0.762 | 0.842 | 0.893 | 0.023* | 0.519 | 0.178 | 0.413 | 0.893 | 0.716 |
| \mathcal{L}_{ind} | 0.909 | 0.304 | 0.230 | 0.112 | 0.187 | 0.134 | 0.848 | 0.702 | 0.478 | 0.644 | 0.884 | 0.932 |
| \mathcal{L}_{cc} | 0.114 | 0.528 | 0.348 | 0.270 | 0.410 | 0.322 | 0.075 | 0.755 | 0.313 | 0.643 | 0.981 | 0.933 |

\mathcal{L}_{ic} : Statistique de couverture inconditionnelle; \mathcal{L}_{ind} : Statistique d'indépendance; \mathcal{L}_{cc} : Statistique de couverture conditionnelle. Le symbole * signifie que le critère est significativement supérieur à la statistique de test au seuil de 5%.

des variances réalisées sont le modèle log-HAR (Corsi *et al.*, 2009, Corsi et Renò, 2012) et le modèle MEM (Engle, 2002). Le modèle MDSV permet aussi de modéliser les variances réalisées. Dans le cadre de ce modèle, la variance réalisée RV_t à la date t vérifie la relation

$$RV_t = V_t \eta_t,$$

où $\{\eta_t\}$ est un processus d'innovations indépendantes et identiquement distribuées selon une loi gamma ($\text{Gamma}(\nu, 1/\nu)$) de moyenne 1 et de paramètre de forme ν et $\{V_t\}$ est le processus MDSV présenté au chapitre 2. Dans cette section, nous effectuons une analyse empirique du modèle MDSV sous sa forme univariée modélisant les variances réalisées. Les performances de trois spécifications (beaucoup de composantes, un nombre moyen de composantes et peu de composantes) du modèle MDSV sont donc comparées à celles de modèles de référence. Les modèles de référence que nous utilisons sont les modèles log-HAR et MEM tels que présentés respectivement aux sections 1.2.4 et 1.2.3, ainsi que le modèle FHMV proposé par Augustyniak *et al.* (2019) avec 10 composantes. Pour prendre en compte l'effet levier dans les modèles log-HAR et MEM, nous utilisons

(1) le modèle log-HAR spécifié comme suit :

$$\begin{cases} \log RV_t &= \beta_0 + \beta_1 \log RV_t^{(1)} + \beta_2 \log RV_t^{(5)} + \beta_3 \log RV_t^{(22)} + \text{lev}_t + \epsilon_t, \\ \text{lev}_t &= \beta_4 r_t^{(1)-} + \beta_5 r_t^{(5)-} + \beta_6 r_t^{(22)-}, \end{cases}$$

où $(\beta_j)_{j=0,1,\dots,6}$ sont des paramètres du modèle, ϵ_t suit une distribution normale de moyenne nulle et de variance σ_ϵ^2 et les variables explicatives sont définies pour $h = 1, 5, 22$ par

$$\log RV_t^{(h)} = \frac{1}{h} \sum_{i=1}^h \log RV_{t-h}$$

et

$$r_t^{(h)-} = \min \left(\frac{1}{h} \sum_{i=1}^h r_{t-h}, 0 \right).$$

Le terme r_t représente le log-rendement à la date t exprimé en pourcentage. Le modèle log-HAR sans effet levier est obtenu en imposant la contrainte $\beta_4 = \beta_5 = \beta_6 = 0$.

(2) le modèle MEM spécifié comme suit :

$$\begin{cases} RV_t &= V_t \eta_t, \\ V_t &= \omega + (1 + \gamma \mathbf{1}_{\{r_{t-1} < 0\}}) RV_{t-1} + \beta V_{t-1}, \end{cases}$$

où ω, γ et β sont des paramètres du modèle et $\{\eta_t\}$ est un processus d'innovations indépendantes et identiquement distribuées selon une loi gamma (**Gamma** $(\nu, 1/\nu)$) de moyenne 1 et de paramètre de forme ν . Le modèle MEM sans effet levier est obtenu en posant $\gamma = 0$.

3.5.1. Estimation du modèle

Le tableau 3.6 présente les résultats (log-vraisemblance, AIC et BIC) obtenus par la méthode du maximum de vraisemblance pour le modèle MDSV et les modèles de référence (log-HAR, MEM et FHMV). Sur chaque ligne du tableau 3.6, les valeurs les plus élevées (désignant le meilleur modèle) sont en gras.

Comme dans le cas des rendements financiers et en accord avec la littérature financière, prendre en compte l'effet levier permet d'améliorer grandement les performances d'ajustement des modèles de variances réalisées. D'après le tableau 3.6, le modèle FHMV est le modèle de référence qui présente les meilleures performance d'ajustement. De façon générale, les performances du modèle MDSV sont similaires à celles du modèle FHMV. En effet, que ce soit dans le cas où l'effet levier est pris en compte ou le cas contraire, lorsque le modèle MDSV ne présente pas la meilleure performance, cette dernière n'est jamais bien loin de celle du meilleur modèle (la différence de performance étant négligeable). Plus particulièrement, dans le cas où on tient compte de l'effet levier, lorsqu'on utilise le critère d'information d'Akaike, le modèle FHMV est meilleur pour modéliser les données des indices S&P 500, FTSE 100 et NIKKEI 225 tandis que le modèle MDSV est préféré aux autres modèles pour modéliser les variances réalisées du NASDAQ 100. En se basant sur le critère d'information bayésien, le modèle MDSV est préféré aux autres modèles pour modéliser le S&P 500 et le NASDAQ 100.

TABLE 3.6 – Comparaison des performances d’ajustement : variances réalisées

| Modèles | Sans effet levier | | | | | | Avec effet levier | | | | | |
|-----------------------------------|----------------------|---------|---------|---------------------|--------------------|---------------------|----------------------|---------|----------------|---------------------|--------------------|---------------------|
| | Modèles de référence | | | MDSV | | | Modèles de référence | | | MDSV | | |
| | log-HAR | MEM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | log-HAR | MEM | FHMV | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| Np | 5 | 4 | 8 | 6 | 6 | 6 | 8 | 5 | 10 | 8 | 8 | 8 |
| S&P 500 ($n = 5016$) | | | | | | | | | | | | |
| \mathcal{L} | -1033.8 | -1416.4 | -968.6 | -1008.1 | -976.4 | -982.9 | -857.6 | -1307.8 | -757.0 | -759.7 | -759.6 | -771.2 |
| AIC | -1038.8 | -1420.4 | -976.6 | -1014.1 | -982.4 | -988.9 | -865.6 | -1312.8 | -767.0 | -767.7 | -767.6 | -779.2 |
| BIC | -1055.1 | -1433.4 | -1002.6 | -1033.6 | -1002.0 | -1008.5 | -891.7 | -1329.1 | -799.6 | -793.8 | -793.6 | -805.3 |
| NASDAQ 100 ($n = 5012$) | | | | | | | | | | | | |
| \mathcal{L} | -1416.7 | -1694.3 | -1409.2 | -1389.9 | -1378.7 | -1406.2 | -1226.1 | -1601.6 | -1210.7 | -1187.2 | -1190.9 | -1201.9 |
| AIC | -1421.7 | -1698.3 | -1417.2 | -1395.9 | -1384.7 | -1412.2 | -1234.1 | -1606.6 | -1220.7 | -1195.2 | -1198.9 | -1209.9 |
| BIC | -1438.0 | -1711.4 | -1443.3 | -1415.5 | -1404.2 | -1431.7 | -1260.2 | -1622.9 | -1253.3 | -1221.3 | -1225.0 | -1236.0 |
| FTSE 100 ($n = 5042$) | | | | | | | | | | | | |
| \mathcal{L} | -1950.1 | -2573.1 | -1814.0 | -1873.2 | -1853.5 | -1874.9 | -1740.5 | -2487.4 | -1592.8 | -1605.3 | -1624.8 | -1636.6 |
| AIC | -1955.1 | -2577.1 | -1822.0 | -1879.2 | -1859.5 | -1880.9 | -1748.5 | -2492.4 | -1602.8 | -1613.3 | -1632.8 | -1644.6 |
| BIC | -1971.4 | -2590.1 | -1848.1 | -1898.8 | -1879.1 | -1900.5 | -1774.6 | -2508.7 | -1635.5 | -1639.4 | -1658.9 | -1670.7 |
| NIKKEI 225 ($n = 4865$) | | | | | | | | | | | | |
| \mathcal{L} | -1542.4 | -2247.6 | -1419.5 | -1424.1 | -1427.1 | -1451.6 | -1420.8 | -2210.3 | -1277.7 | -1288.9 | -1303.9 | -1330.1 |
| AIC | -1547.4 | -2251.6 | -1427.5 | -1430.1 | -1433.1 | -1457.6 | -1428.8 | -2215.3 | -1287.7 | -1296.9 | -1311.9 | -1338.1 |
| BIC | -1563.7 | -2264.5 | -1453.5 | -1449.6 | -1452.6 | -1477.0 | -1454.8 | -2231.5 | -1320.1 | -1322.9 | -1337.9 | -1364.1 |

Np : Nombre de paramètres, \mathcal{L} : log-vraisemblance, AIC : Critère d’information d’Akaike, BIC : Critère d’information bayésien, n : Taille d’échantillon. Les valeurs les plus élevées sont en gras.

Lorsqu’on ne tient pas compte de l’effet levier, parmi les spécifications du modèle MDSV, celle avec un nombre moyen de composantes ($N = 6, K = 3$) présente la meilleure performance d’ajustement sur les indices financiers S&P 500, NASDAQ 100 et FTSE 100. Pour le NIKKEI 225, la meilleure spécification est celle avec peu de composantes ($N = 3, K = 10$). Par contre, lorsqu’on tient compte de l’effet levier dans la modélisation, la spécification ayant la meilleure performance d’ajustement pour modéliser les variances réalisées des indices financiers NASDAQ 100, FTSE 100 et NIKKEI 225 est celle avec peu de composantes ($N = 3, K = 10$). Pour le S&P 500, la spécification avec un nombre moyen de composantes ($N = 6, K = 3$) présente les meilleures performances d’ajustement.

3.5.2. Performance de prévision du modèle

Le modèle MDSV univarié modélisant les variances réalisées permet aussi de produire des prévisions de volatilité. Le tableau 3.7 présente les écarts de prévision quadratique et absolu moyens, ainsi que la quasi-vraisemblance des modèles de référence et du modèle MDSV pour les horizons $h = 1, 5, 25, 100$ jours. Pour chaque indice, chaque critère et chaque horizon (h), les valeurs les plus faibles (désignant la meilleure

performance de prévision de volatilité) sont en gras. La signification statistique des différences entre les critères est évaluée par le test de Diebold-Mariano (Diebold et Mariano, 2002). Pour ce test, nous utilisons le modèle FHMV comme modèle de référence. Notons également que les prévisions de rendements sont nécessaires pour produire des prévisions de variances réalisées d’horizon supérieur à 1 jour pour les modèles qui incluent un effet de levier. À cette fin, nous supposons que les futurs pourcentages de rendements financiers sont indépendants et identiquement distribués selon une loi normale centrée réduite. La prévision de l’effet de levier pourrait être améliorée en considérant un modèle joint pour les rendements et les variances réalisées ; ce qui fait l’objet de la section 3.6.

Contrairement au cas des rendements financiers, le tableau 3.7 indique que l’impact de l’effet levier est négligeable et son effet sur les critères de prévision semble aléatoire (surtout à long terme). Cependant, en se basant sur le QLIK à l’horizon d’un jour, on constate que le fait de tenir compte de l’effet levier permet d’améliorer légèrement les prévisions de tous les modèles, sauf le modèle log-HAR. De façon générale, le modèle MDSV permet de faire de meilleures prévisions à long terme (horizon h élevé) par rapport aux autres modèles. Par exemple, le modèle MDSV présente la plus petite erreur de prévision (RMSFE, MAFE et QLIK) pour tous les indices financiers à l’horizon de 100 jours. De plus, dans 37 cas sur 48 (horizon de plus d’une semaine), le modèle MDSV présente une erreur de prévision (RMSFE, MAFE et QLIK) significativement (1%, 5% et 10%) plus faible que celle du modèle FHMV. Plus particulièrement, sur les données du S&P 500 et du NASDAQ 100, le modèle MDSV est préféré aux modèles de référence aussi bien à court qu’à long terme (RMSFE et MAFE). Sur les données du FTSE 100 et du NIKKEI 225, le modèle MDSV est préféré à un horizon supérieur à une semaine (RMSFE, QLIK) alors que le modèle FHMV est préféré à l’horizon d’un jour (QLIK). En se basant sur le MAFE, le test de Diebold-Mariano indique le modèle MDSV domine significativement (1%, 5% et 10%) le modèle FHMV à long terme (horizon d’une semaine ou plus) sur les données des quatre indices financiers.

Comme dans le cas de la modélisation des rendements financiers, nous avons étudié trois spécifications du modèle MDSV : la spécification avec peu de composantes ($N = 3, K = 10$), la spécification avec un nombre moyen de composantes ($N = 6, K = 3$) et celle avec beaucoup de composantes ($N = 10, K = 2$). Ces trois spécifications ont des performances de prévision similaires aussi bien à court qu’à long terme. Cependant, la spécification avec beaucoup de composantes ($N = 10, K = 2$) est celle qui présente, le plus souvent, les valeurs de RMSFE, MAFE et QLIK les plus faibles. De façon similaire au cas de la modélisation des rendements financiers, ces spécifications présentent aussi une interprétation économique intéressante. Les composantes sont interprétées comme de nouvelles informations qui affectent la volatilité et dont le degré de persistance est décroissant. Ainsi, la première composante influence beaucoup plus longuement la volatilité que la seconde composante, et ainsi de suite.

3.6. Modélisation conjointe : rendements financiers et variances réalisées

Les rendements financiers et les variances réalisées sont les outils les plus utilisés pour estimer et prévoir la volatilité des indices financiers. Des études récentes (Takahashi *et al.*, 2009, Hansen *et al.*, 2012, Hansen et Huang, 2016, Liu et Maheu, 2018, Takahashi *et al.*, 2021) montrent que le fait de modéliser simultanément les deux mesures permet de mieux capter le niveau actuel de volatilité, une information par la suite utilisée pour former des anticipations sur la volatilité de la prochaine période. Le modèle MDSV présente une forme jointe permettant ainsi de modéliser simultanément les rendements financiers et les variances réalisées. Dans le cadre de cette thèse, dans un modèle joint, le rendement financier r_t et la variance réalisée RV_t à la date t vérifient les relations

$$r_t = \sqrt{V_t} \epsilon_t, \quad (3.6.1)$$

$$\log RV_t = \xi + \varphi \log V_t + \delta_1 \epsilon_t + \delta_2 (\epsilon_t^2 - 1) + \gamma \varepsilon_t, \quad (3.6.2)$$

où $\xi, \varphi, \delta_1, \delta_2 \in \mathbb{R}$ et $\gamma \in (0, \infty)$ sont des paramètres, et $\{\epsilon_t\}$ et $\{\varepsilon_t\}$ sont des processus d'innovation mutuellement indépendants et identiquement distribués selon une loi normale de moyenne 0 et de variance 1. Cette section concerne une analyse empirique du modèle MDSV sous sa forme jointe modélisant simultanément les rendements financiers et les variances réalisées. Nous comparons ainsi les performances d'ajustement, de prévision et de gestion du risque des trois spécifications (peu de composantes, un nombre moyen de composantes et beaucoup de composantes) à celles de modèles de référence. Nous utilisons comme modèles de référence deux modèles joints financiers populaires : le modèle Real EGARCH (*Realized Exponential Generalized Autoregressive Conditional Heteroscedasticity*) de Hansen et Huang (2016) et le modèle modèle MS-RV (modèle joint de variances réalisées à chaîne de Markov cachée), une modification appropriée du modèle proposé par Liu et Maheu (2018). Ces modèles de référence relient également les rendements financiers et les variances réalisées par les équations (3.6.1) et (3.6.2) comme le modèle MDSV. La différence fondamentale entre ces modèles est la manière par laquelle le processus de volatilité $V = \{V_t\}$ est modélisé.

Tandis que le modèle MDSV modélise le processus de volatilité V par le processus MDSV présenté au chapitre 2, le modèle Real EGARCH de Hansen et Huang (2016) stipule que le processus de volatilité V vérifie la relation

$$\log V_t = \omega + \beta \log V_{t-1} + \tau_1 \epsilon_t + \tau_2 (\epsilon_t^2 - 1) + \gamma \eta \varepsilon_{t-1},$$

où $\omega, \beta, \tau_1, \tau_2 \in \mathbb{R}$, et $\eta \in (0, \infty)$ sont des paramètres. L'effet levier est capté par le processus $\{\tau_1 \epsilon_t + \tau_2 (\epsilon_t^2 - 1)\}$ et le modèle Real EGARCH sans effet levier est obtenu en imposant qu'il soit nul, c'est-à-dire $\tau_1 = \tau_2 = 0$.

Dans le cas du modèle MS-RV, le processus V est un lié à une chaîne de Markov $\{s_t\}$ à espace d'états $\{1, \dots, K\}$. Ainsi, le processus de volatilité vérifie l'équation

$$V_t = \sigma_{s_t}^2,$$

où $\sigma_1, \sigma_2, \dots, \sigma_K$ sont des paramètres à estimer. Le modèle MS-RV que nous présentons s'inspire du modèle de Liu et Maheu (2018)⁴. Un fait intéressant avec le modèle MS-RV est qu'il nous permet de comparer un modèle factoriel à chaîne de Markov cachée (le modèle MDSV) avec un modèle à chaîne de Markov classique (le modèle MS-RV). Cependant, le modèle MS-RV ne prend pas en compte l'effet levier. Pour le comparer aux modèles intégrant un effet levier, nous multiplions la volatilité par le processus effet levier multiplicatif L_t présenté à la section 2.5 :

$$V_t = \sigma_{s_t}^2 L_t.$$

Dans ce qui suit, nous utilisons le modèle MS-RV avec $K = 2$ et $K = 4$ états.

Dans les modèles joints présentés dans cette thèse c'est-à-dire, les modèles qui relient les rendements financiers et les variances réalisées par les relations (3.6.1) et (3.6.2), les distributions conditionnelles des rendements financiers et des variances réalisées sont :

$$\begin{cases} r_t | V_t \sim \text{Normal}(0, V_t), \\ \log(RV_t) | V_t, r_t \sim \text{Normal}\left(\xi + \varphi \log V_t + \delta_1 \left(\frac{r_t}{\sqrt{V_t}}\right) + \delta_2 \left(\left(\frac{r_t}{\sqrt{V_t}}\right)^2 - 1\right), \gamma^2\right). \end{cases}$$

3.6.1. Estimation du modèle

Le tableau 3.8 présente les résultats (log-vraisemblance, AIC et BIC) obtenus par la méthode du maximum de vraisemblance pour le modèle MDSV et les modèles de référence (Real EGARCH, MS-RV). A titre indicatif, les résultats du modèle de Liu et Maheu (2018) sont présentés à l'annexe B.1. Sur chaque ligne du tableau 3.8, les valeurs les plus élevées (désignant le meilleur modèle) sont en gras.

Les résultats du tableau 3.8 concordent avec la théorie financière : prendre en compte l'effet levier permet d'améliorer les performances d'ajustement de tous les modèles. En tenant compte de l'effet levier ou non, le modèle MDSV présente une meilleure performance d'ajustement que les modèles de référence. Lorsqu'on ne tient pas compte de l'effet levier, la spécification du modèle MDSV avec un nombre moyen de composantes ($N = 6, K = 3$) présente les meilleures performances d'ajustement des données des indices S&P 500, NASDAQ 100 et FTSE 100, alors que la spécification avec peu de composantes ($N = 3, K = 10$) est préférée sur les données du NIKKEI 225. Par contre, en tenant compte de l'effet levier, on observe que la spécification ($N = 3, K = 10$) est celle qui offre les meilleures performances pour tous les indices financiers.

Afin de mesurer la robustesse des résultats empiriques obtenus sur les quatre séries de données (S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225), une analyse similaire est effectuée sur les 27 autres indices financiers disponibles sur le site web de *Oxford-Man Institute Quantitative Finance*. Le graphique 3.6 présente les différences entre les BIC obtenus avec les trois spécifications MDSV et ceux obtenus avec le modèle Real EGARCH en prenant compte de l'effet levier.

De façon similaire aux résultats du tableau 3.8, ce graphique indique que le modèle MDSV présente les meilleures performances d'ajustement pour presque tous les indices financiers comparativement au modèle

4. Le modèle MS-RV présenté dans cette thèse donne de meilleurs résultats empiriques que le modèle de Liu et Maheu (2018) dont il s'inspire (voir les tableaux B.1, B.2 et B.3).

TABLE 3.8 – Comparaison des performances d’ajustement : modèle joint

| Modèles | Sans effet levier | | | | | | Avec effet levier | | | | | |
|-----------------------------------|----------------------|------------------|------------------|---------------------|--------------------|---------------------|----------------------|------------------|------------------|---------------------|--------------------|---------------------|
| | Modèles de référence | | | MDSV | | | Modèles de référence | | | MDSV | | |
| | Real EGARCH | MS-RV $K = 2$ | MS-RV $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | Real EGARCH | MS-RV $K = 2$ | MS-RV $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| Np | 8 | 5 | 17 | 10 | 10 | 10 | 10 | 7 | 19 | 12 | 12 | 12 |
| S&P 500 ($n = 5016$) | | | | | | | | | | | | |
| \mathcal{L} | -7342.6 | -8898.2 | -7616.6 | -6979.0 | -6977.0 | -6996.0 | -6786.1 | -7765.8 | -6892.6 | -6673.8 | -6687.9 | -6701.0 |
| AIC | -7350.6 | -8907.2 | -7637.6 | -6989.0 | -6987.0 | -7006.0 | -6796.1 | -7776.8 | -6915.6 | -6685.8 | -6699.9 | -6713.0 |
| BIC | -7376.7 | -8936.5 | -7706.1 | -7021.6 | -7019.6 | -7038.6 | -6828.7 | -7812.6 | -6990.6 | -6724.9 | -6739.0 | -6752.2 |
| NASDAQ 100 ($n = 5012$) | | | | | | | | | | | | |
| \mathcal{L} | -9172.1 | -10798.2 | -9999.3 | -8896.5 | -8892.0 | -8914.5 | -8661.3 | -9787.5 | -8965.5 | -8546.9 | -8568.3 | -8590.3 |
| AIC | -9180.1 | -10807.2 | -10020.3 | -8906.5 | -8902.0 | -8924.5 | -8671.3 | -9798.5 | -8988.5 | -8558.9 | -8580.3 | -8602.3 |
| BIC | -9206.2 | -10836.5 | -10088.7 | -8939.1 | -8934.6 | -8957.1 | -8703.9 | -9834.3 | -9063.4 | -8598.0 | -8619.4 | -8641.5 |
| FTSE 100 ($n = 5042$) | | | | | | | | | | | | |
| \mathcal{L} | -8282.8 | -9435.6 | -8356.9 | -8044.7 | -8037.3 | -8050.7 | -7809.9 | -8388.7 | -7902.9 | -7761.4 | -7804.8 | -7789.4 |
| AIC | -8290.8 | -9444.6 | -8377.9 | -8054.7 | -8047.3 | -8060.7 | -7819.9 | -8399.7 | -7925.9 | -7773.4 | -7816.8 | -7801.4 |
| BIC | -8316.9 | -9474.0 | -8446.4 | -8087.3 | -8079.9 | -8093.3 | -7852.5 | -8435.6 | -8001.0 | -7812.6 | -7855.9 | -7840.6 |
| NIKKEI 225 ($n = 4865$) | | | | | | | | | | | | |
| \mathcal{L} | -9411.3 | -10612.8 | -9871.3 | -9059.8 | -9076.1 | -9122.9 | -9004.3 | -9656.4 | -9028.5 | -8839.2 | -8863.2 | -8877.5 |
| AIC | -9419.3 | -10621.8 | -9892.3 | -9069.8 | -9086.1 | -9132.9 | -9014.3 | -9667.4 | -9051.5 | -8851.2 | -8875.2 | -8889.5 |
| BIC | -9445.2 | -10651.0 | -9960.4 | -9102.3 | -9118.6 | -9165.3 | -9046.8 | -9703.1 | -9126.1 | -8890.1 | -8914.1 | -8928.4 |

Np : Nombre de paramètres, \mathcal{L} : log-vraisemblance, AIC : Critère d’information d’Akaike, BIC : Critère d’information bayésien, n : Taille d’échantillon. Les valeurs les plus élevées sont en gras.

Real EGARCH. La différence de BIC entre les modèles MDSV et Real EGARCH n’est pas constante. Alors que cette différence est de l’ordre de 100 pour les indices S&P 500 et NASDAQ 100, elle prend des valeurs proches de 800 pour l’indice Karachi SE 100 (KSE). Le graphique 3.6 nous indique également que la spécification ($N = 3, K = 10$) n’est pas celle qui présente la BIC la plus grande parmi les trois spécifications pour toutes les indices. Cependant, elle présente la BIC la plus grande pour 22 indices financiers sur 31.

3.6.2. Performance de prévision du modèle

La représentation jointe présente un avantage supplémentaire en terme de prévision. En effet, à travers un modèle joint, il est possible de faire de la prévision de volatilité en utilisant les carrés de rendements financiers et/ou les variances réalisées. Ainsi, outre la log-vraisemblance prédictive, nous pouvons analyser aussi bien la capacité du modèle MDSV joint à bien prédire les variances réalisées que celle à bien prédire le carré des rendements financiers. Le tableau 3.9 présente la log-vraisemblance prédictive pour les trois spécifications du modèle MDSV et pour les modèles de référence (Real EGARCH, MS-RV). Pour chaque indice, les valeurs les plus élevées (désignant le meilleur modèle) sont en gras.

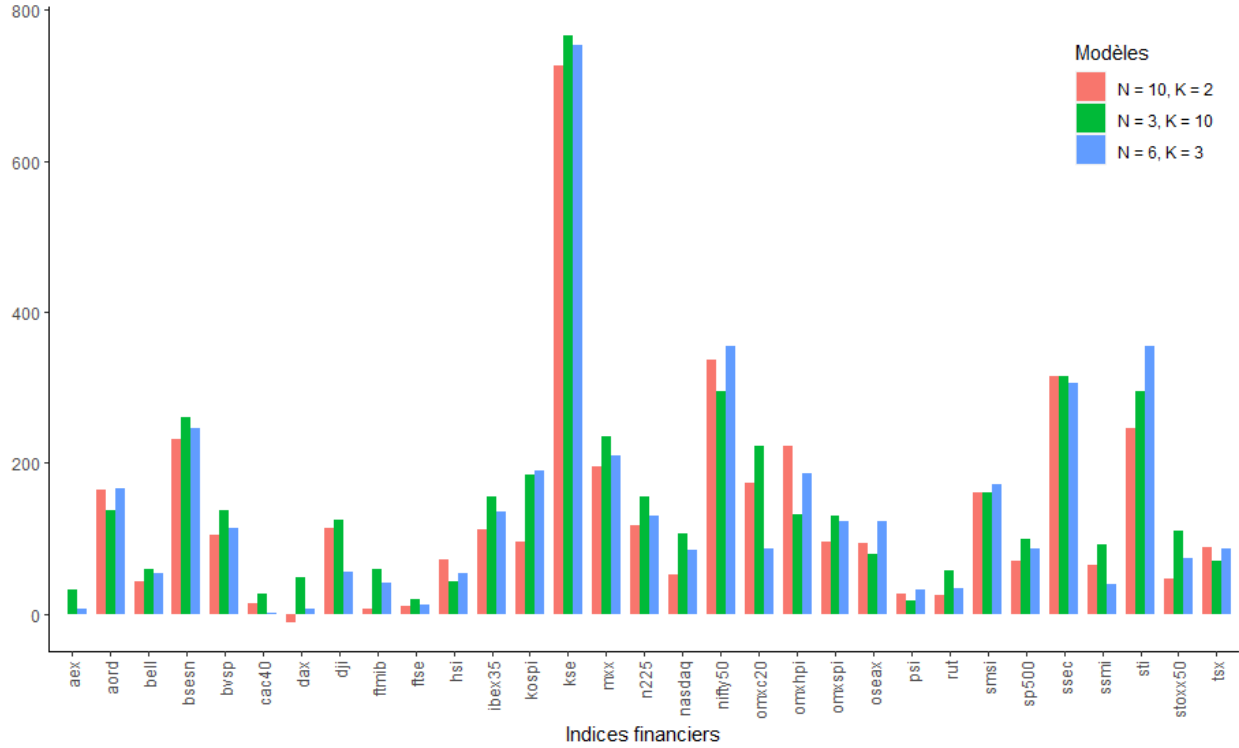


FIGURE 3.6 – Différence entre le BIC des trois spécifications du modèle MDSV et le modèle Real EGARCH en tenant compte de l’effet levier

TABLE 3.9 – Log-vraisemblance prédictive : modèle joint

| Modèles | Sans effet levier | | | | | | Avec effet levier | | | | | |
|------------|----------------------|------------------|------------------|---------------------|--------------------|---------------------|----------------------|------------------|------------------|---------------------|--------------------|---------------------|
| | Modèles de référence | | | MDSV | | | Modèles de référence | | | MDSV | | |
| | Real EGARCH | MS-RV $K = 2$ | MS-RV $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | Real EGARCH | MS-RV $K = 2$ | MS-RV $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| S&P 500 | -1124.7 | -777.0 | -740.6 | -711.5 | -713.2 | -718.3 | -1113.2 | -754.0 | -718.0 | -708.0 | -711.2 | -712.7 |
| NASDAQ 100 | -1082.8 | -990.6 | -943.3 | -924.5 | -924.2 | -923.2 | -1069.4 | -963.0 | -931.4 | -924.1 | -923.4 | -924.2 |
| FTSE 100 | -1111.1 | -808.8 | -796.0 | -790.5 | -792.0 | -792.7 | -1090.9 | -795.0 | -787.3 | -782.4 | -789.5 | -783.8 |
| NIKKEI 225 | -1059.8 | -1031.3 | -986.9 | -975.5 | -978.0 | -978.7 | -1055.1 | -998.1 | -981.1 | -973.7 | -975.1 | -975.4 |

Les valeurs les plus élevées sont en gras.

La log-vraisemblance prédictive mesure à quel point le rendement du lendemain est probable compte tenu de l’information disponible jusqu’à présent. D’après le tableau 3.9, prendre en compte l’effet levier permet d’améliorer les chances de faire de bonnes prévisions puisque les densités prédictives cumulatives de presque tous les modèles augmentent lorsqu’un effet levier est incorporé (23 cas sur 24). Cependant, la différence de performances par rapport au fait d’ignorer l’effet levier n’est pas très grande, indiquant qu’un modèle sans effet levier pourrait produire de bonnes prévisions. D’après ce tableau, les performances de

prévision de densité du modèle MDSV sont meilleures que celles des modèles de référence. De façon générale, les performances de prévision de densité des différentes spécifications du modèle MDSV sont très similaires. Dans cette étude, la spécification avec peu de composantes ($N = 3, K = 10$), incluant l'effet levier, présente les valeurs les plus élevées de log-vraisemblance prédictive pour les indices S&P 500, le FTSE 100 et NIKKEI 225. Pour le NASDAQ 100, la spécification avec beaucoup de composantes ($N = 10, K = 2$), sans l'effet levier, présente la meilleure performance de prévision de densité cumulative.

Le tableau 3.10 présente les écarts de prévision quadratique et absolu, ainsi que la quasi-vraisemblance des modèles de référence et du modèle MDSV pour les horizons $h = 1, 5, 25, 100$ jours. Les mesures de performance présentées dans ce tableau se basent sur les prévisions de variances réalisées obtenues grâce à chacun des modèles. En effet, les modèles joints permettent la prévision aussi bien des rendements financiers que des variances réalisées. Nous présentons ici les résultats pour les prévisions de variances réalisées. Cependant, on parvient à des conclusions similaires lorsqu'on calcule les mesures de performances sur les prévisions de rendements financiers. Dans le tableau 3.10, les valeurs les plus faibles (désignant le meilleur modèle) pour chaque indice et à chaque horizon (h) sont en gras. La signification statistique des différences de critères est évaluée par le test de Diebold-Mariano. Pour ce test, nous utilisons le modèle Real EGARCH comme modèle de référence.

D'après le tableau 3.10, on constate que l'effet levier permet d'améliorer les performances de prévision à court terme de presque tous les modèles (188 cas sur 192). L'introduction de l'effet levier permet aussi d'améliorer très légèrement les performances de prévisions à long terme des modèles Real EGARCH et MDSV. De façon générale, le modèle MDSV permet de faire de meilleures prévisions que les modèles de référence, surtout à long terme. Par exemple, en se basant sur le critère MAFE, le modèle MDSV est préféré aux autres modèles pour tous les quatre indices financiers et à tous les horizons de prévision. Dans le cas du RMSFE, le modèle MDSV joint est préféré pour prédire les variances réalisées des indices financiers S&P 500 et NASDAQ 100 à tous les horizons. Ce modèle est également préféré pour les prévisions d'horizons de plus d'un jour sur les deux autres indices (FTSE 100 et NIKKEI 225). Le critère RMSFE du modèle MDSV à l'horizon d'un jour est cependant similaire à celui du meilleur modèle (le modèle Real EGARCH). En se basant sur le QLIK, le modèle MDSV est préféré pour la prévision des données financières des indices S&P 500 et NASDAQ 100 à tous les horizons. En effet, sur les 48 cas testés, seulement 4 d'entre eux ne sélectionnent pas le modèle MDSV comme optimal. Notons que le test de Diebold-Mariano montre que le modèle MDSV produit des RMSFE et MAFE significativement inférieurs à ceux du modèle Real EGARCH pour des horizons supérieurs à une semaine sur toutes les données utilisées. De plus, parmi les trois spécifications du modèle MDSV, celle avec beaucoup de composantes ($N = 10, K = 2$) est celle qui présente les meilleures performances de prévision la plupart du temps (30 cas sur 48).

Remarquons que le modèle MS-RV à K états peut être vu comme un modèle MDSV avec une composante et K états puisque les deux modèles partagent une structure de chaîne de Markov. On constate que les performances de prévision du modèle MS-RV avec 4 états sont meilleures que celles du modèle MS-RV avec 2

états. Cependant, le modèle MDSV domine toujours le modèle MS-RV. Le modèle MDSV apporte, en effet, un nombre d'états suffisamment grand pour avoir une meilleure performance de prévision que le modèle MS-RV.

3.6.3. Performance de gestion du risque du modèle

Un autre atout des modèles joints est qu'ils permettent d'évaluer le risque dans les rendements financiers, contrairement aux modèles de variances réalisées. Le tableau 3.11 présente les statistiques des tests de couverture inconditionnelle, d'indépendance et de couverture conditionnelle basées sur la valeur à risque conditionnelle de niveau 95% pour le modèle MDSV et les modèles de référence. Lorsque le critère est significativement supérieur à la statistique de test au seuil de 5%, la valeur-p est marquée du symbole *.

TABLE 3.11 – Performance de gestion du risque : rendements financiers et variances réalisées

| Modèles | Sans effet levier | | | | | | Avec effet levier | | | | | |
|---------------------|----------------------|------------------|------------------|---------------------|--------------------|---------------------|----------------------|------------------|------------------|---------------------|--------------------|---------------------|
| | Modèles de référence | | | MDSV | | | Modèles de référence | | | MDSV | | |
| | Real EGARCH | MS-RV $K = 2$ | MS-RV $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | Real EGARCH | MS-RV $K = 2$ | MS-RV $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| S&P 500 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.185 | 0.087 | 0.491 | 0.073 | 0.052 | 0.052 | 0.313 | 0.178 | 0.636 | 0.007* | 0.025* | 0.036* |
| \mathcal{L}_{ind} | 0.201 | 0.380 | 0.289 | 0.633 | 0.691 | 0.691 | 0.776 | 0.478 | 0.585 | 0.575 | 0.334 | 0.371 |
| \mathcal{L}_{cc} | 0.183 | 0.157 | 0.449 | 0.179 | 0.140 | 0.140 | 0.577 | 0.313 | 0.770 | 0.023* | 0.050 | 0.074 |
| NASDAQ 100 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.842 | 0.519 | 0.598 | 0.395 | 0.313 | 0.313 | 0.716 | 0.321 | 0.636 | 0.313 | 0.598 | 0.395 |
| \mathcal{L}_{ind} | 0.490 | 0.265 | 0.872 | 0.756 | 0.776 | 0.700 | 0.932 | 0.586 | 0.762 | 0.700 | 0.336 | 0.756 |
| \mathcal{L}_{cc} | 0.773 | 0.437 | 0.859 | 0.664 | 0.577 | 0.558 | 0.933 | 0.527 | 0.854 | 0.558 | 0.547 | 0.664 |
| FTSE 100 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.185 | 0.973 | 0.102 | 0.102 | 0.052 | 0.052 | 0.973 | 0.126 | 0.636 | 0.313 | 0.185 | 0.313 |
| \mathcal{L}_{ind} | 0.001* | 0.002* | 0.105 | 0.010* | 0.016* | 0.016* | 0.440 | 0.118 | 0.304 | 0.373 | 0.201 | 0.373 |
| \mathcal{L}_{cc} | 0.002* | 0.008* | 0.070 | 0.009* | 0.008* | 0.008* | 0.741 | 0.092 | 0.528 | 0.404 | 0.183 | 0.404 |
| NIKKEI 225 | | | | | | | | | | | | |
| \mathcal{L}_{ic} | 0.762 | 0.178 | 0.519 | 0.138 | 0.185 | 0.395 | 0.519 | 0.058 | 0.519 | 0.491 | 0.491 | 0.598 |
| \mathcal{L}_{ind} | 0.346 | 0.478 | 0.076 | 0.495 | 0.469 | 0.329 | 0.702 | 0.970 | 0.702 | 0.289 | 0.657 | 0.599 |
| \mathcal{L}_{cc} | 0.613 | 0.313 | 0.169 | 0.264 | 0.320 | 0.433 | 0.755 | 0.166 | 0.755 | 0.449 | 0.714 | 0.758 |

\mathcal{L}_{ic} : Statistique de couverture inconditionnelle ; \mathcal{L}_{ind} : Statistique d'indépendance ; \mathcal{L}_{cc} : Statistique de couverture conditionnelle. Le symbole * signifie que le critère est significativement supérieur à la statistique de test au seuil de 5%.

D'après le tableau 3.11, les trois tests (couverture inconditionnelle, indépendance et couverture conditionnelle) effectués sur les prévisions de valeurs à risque à 95% par les trois spécifications du modèle MDSV (tenant compte de l'effet levier) ne permettent pas de rejeter hypothèse nulle au seuil de 5% pour les données financières des indices NASDAQ 100, FTSE 100 et NIKKEI 225. La valeur à risque de niveau 95% est donc

valide pour le modèle MDSV sur ces trois séries de données financières. Sur la série de l'indice S&P 500, le test de couverture conditionnelle pour un nombre moyen ou élevé de composantes ne rejette pas l'hypothèse nulle, mais la rejette lorsqu'il y a peu de composantes ($N = 3, K = 10$). De façon générale, le modèle MDSV sous sa forme jointe produit des valeurs à risque de niveau 95% valides. Ainsi, en utilisant le modèle MDSV, on prévoit que les pertes quotidiennes dépasseront la valeur à risque prédite environ 5% du temps.

3.7. MDSV : forme jointe ou forme univariée ?

Une question légitime à se poser est de savoir s'il est plus adapté de faire une modélisation univariée des rendements financiers et/ou des variances réalisées, ou plutôt de modéliser ces deux processus conjointement. En d'autres mots, un modèle joint a-t-il réellement un avantage comparativement à un modèle univarié ? Le modèle MDSV présente l'avantage de disposer à la fois d'une forme jointe et d'une forme univariée. Cela nous permet ainsi de comparer convenablement les deux formes et donc de répondre à cette question. Dans cette section, nous comparons les performances d'ajustement, de prévision et de gestion du risque de la forme jointe du modèle MDSV à celle de sa forme univariée modélisant les rendements financiers ou les variances réalisées. Pour cette analyse, l'effet levier est pris en compte dans tous les modèles.

3.7.1. Ajustement et prévision des rendements financiers

Afin de comparer la performance d'ajustement d'un modèle joint à celui d'un modèle univarié modélisant les rendements financiers, nous utilisons la log-vraisemblance intégrée que nous désignons par \mathcal{L}_r . En effet, la log-vraisemblance intégrée peut être calculée dans les modèles joints en remarquant que

$$\mathcal{L} = \underbrace{\log p_{\theta}(r_1) + \sum_{t=2}^T \log p_{\theta}(r_t | \mathcal{F}_{t-1})}_{\mathcal{L}_r} + \underbrace{\log p_{\theta}(RV_1 | r_1) + \sum_{t=2}^T \log p_{\theta}(RV_t | r_t, \mathcal{F}_{t-1})}_{\mathcal{L}_{RV|r}},$$

où \mathcal{L} désigne la log-vraisemblance du modèle joint et \mathcal{F}_{t-1} désigne toute l'information disponible jusqu'à la date $t - 1$. La log-vraisemblance intégrée d'un modèle joint est ainsi comparée à la log-vraisemblance d'un modèle univarié de rendements financiers pour vérifier s'il est plus adéquat de modéliser les rendements financiers avec un modèle joint ou avec un modèle univarié. Contrairement à la log-vraisemblance d'un modèle univarié, le terme \mathcal{F}_{t-1} dans la log-vraisemblance intégrée incorpore l'information disponible sur les variances réalisées. Cette information pourrait donc permettre de mieux comprendre la distribution des rendements financiers.

Le tableau 3.12 présente la log-vraisemblance des trois spécifications du modèle MDSV sous leur forme univariée modélisant les rendements financiers ainsi que la log-vraisemblance intégrée des spécifications correspondantes du modèle MDSV joint. Le tableau 3.12 présente aussi la log-vraisemblance prédictive des trois spécifications du modèle MDSV sous les deux formes (jointe et univariée). Dans ce tableau, les valeurs des log-vraisemblances et des densités prédictives les plus élevées (par spécification) sont en gras.

TABLE 3.12 – Log-vraisemblance et log-vraisemblance prédictive : forme univariée ou forme jointe

| (N, K) | Log-vraisemblance | | | | | | log-vraisemblance prédictive | | | | | |
|------------|-------------------|----------------|---------|----------------|---------|----------------|------------------------------|---------------|--------|---------------|---------------|---------------|
| | (3, 10) | | (6, 3) | | (10, 2) | | (3, 10) | | (6, 3) | | (10, 2) | |
| Forme | univ | joint | univ | joint | univ | joint | univ | joint | univ | joint | univ | joint |
| S&P 500 | -6563.7 | -6300.6 | -6606.7 | -6326.5 | -6584.5 | -6329.2 | -735.7 | -708.0 | -734.5 | -711.2 | -729.1 | -712.7 |
| NASDAQ 100 | -7950.0 | -7688.6 | -7947.5 | -7688.1 | -7947.6 | -7691.8 | -951.3 | -924.1 | -950.4 | -923.4 | -950.6 | -924.2 |
| FTSE 100 | -6659.0 | -6471.5 | -6657.8 | -6498.7 | -6657.9 | -6476.8 | -796.2 | -782.4 | -796.6 | -789.5 | -795.6 | -783.2 |
| NIKKEI 225 | -8061.3 | -7863.4 | -8051.0 | -7876.8 | -8052.0 | -7885.6 | -983.9 | -973.7 | -976.5 | -975.1 | -972.8 | -975.4 |

univ : forme univariée, joint : forme jointe. Les valeurs les plus élevées sont en gras.

D'après le tableau 3.12, la log-vraisemblance intégrée de la forme jointe est plus élevée que la log-vraisemblance de la forme univariée pour toutes les spécifications et pour tous les indices financiers. Ce résultat confirme ainsi que la forme jointe améliore grandement l'ajustement du modèle aux rendements financiers par rapport à la forme univariée. En effet, l'information disponible sur les variances réalisées permet de mieux comprendre la distribution des rendements financiers. Plus techniquement, le filtrage de la volatilité dans un modèle joint utilise l'information disponible sur les variances réalisées. Cette information n'est pas disponible dans un modèle univarié de rendements financiers. Ceci explique pourquoi la log-vraisemblance intégré de la forme jointe est plus élevée que la log-vraisemblance de la forme univariée. L'analyse de la log-vraisemblance prédictive montre aussi que la forme jointe est préférable à la forme univariée. En effet, dans tous les cas à l'exception la spécification avec beaucoup de composantes ($N = 10, K = 2$) ajustée au NIKKEI 225, la forme jointe produit une log-vraisemblance prédictive plus élevée que la forme univariée. Cela indique alors que la forme jointe produira vraisemblablement de meilleures prédictions des rendements financiers que la forme univariée.

Le tableau 3.13 présente les trois autres mesures de performance des prévisions (les écarts quadratique et absolu moyens et la quasi-vraisemblance) pour différents horizons, basées sur la prévision du carré des rendements financiers. Pour chaque indice financier, chaque mesure de performance et chaque horizon, les valeurs les plus faibles sont en gras. La signification statistique des différences de critères (entre la forme jointe et la forme univariée) est évaluée par le test de Diebold-Mariano. Pour ce test, la forme univariée est considérée comme modèle de référence.

D'après le tableau 3.13, la forme jointe semble produire des meilleures performances de prévision. En effet, dans 123 cas sur 144, la forme jointe présente une erreur de prévision (RMSFE, MAFE ou QLIK) plus faible que la forme univariée. Par exemple, en se basant sur les trois critères de performance (RMSFE, MAFE et QLIK), la forme jointe de chacune des spécifications a les valeurs les plus faibles pour les indices FTSE 100 et NIKKEI 225. De plus, le test de Diebold-Mariano indique que la forme jointe produit des MAFE significativement (5% et 1%) plus faibles, à tous les horizons, que ceux de la forme univariée sur les données des indices financiers FTSE 100 et NIKKEI 225.

TABLE 3.13 – Prédiction du carré des rendements financiers : forme univariée ou forme jointe

| | | RMSFE | | | | MAFE | | | | QLIK | | | |
|--------------------|-----------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| horizon(h) | | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 |
| S&P 500 | | | | | | | | | | | | | |
| MDSV (3, 10) | univariée | 1.62 | 1.03 | 1.05 | 1.36 | 0.75 | 0.54 | 0.63 | 0.93 | 0.16 | 1.87 | 3.76 | 5.44 |
| | jointe | 1.63 | 0.87[†] | 0.68[†] | 0.55[‡] | 0.66[‡] | 0.41[‡] | 0.43[‡] | 0.45[‡] | 0.12 | 1.82 | 3.73 | 5.31[‡] |
| MDSV (6, 3) | univariée | 1.59 | 0.92 | 0.74 | 0.72 | 0.70 | 0.48 | 0.49 | 0.55 | 0.16 | 1.88 | 3.75 | 5.40 |
| | jointe | 1.59 | 0.86[*] | 0.68 | 0.57 | 0.65[‡] | 0.40[‡] | 0.40[†] | 0.41[‡] | 0.10[*] | 1.84 | 3.83 | 5.58 |
| MDSV (10, 2) | univariée | 1.61 | 0.89 | 0.68 | 0.54 | 0.70 | 0.46 | 0.49 | 0.45 | 0.15 | 1.86 | 3.72 | 5.31 |
| | jointe | 1.58 | 0.86[*] | 0.69 | 0.58 | 0.64[‡] | 0.40[‡] | 0.39[†] | 0.42 | 0.10[*] | 1.85 | 3.87 | 5.68 |
| NASDAQ 100 | | | | | | | | | | | | | |
| MDSV (3, 10) | univariée | 2.08 | 1.15 | 1.00 | 1.23 | 1.09 | 0.73 | 0.72 | 0.95 | 0.73 | 2.41 | 4.19 | 5.80 |
| | jointe | 2.16 | 1.23 | 0.94 | 0.86[‡] | 1.08 | 0.70 | 0.67[*] | 0.70[‡] | 0.66[†] | 2.35[†] | 4.16 | 5.73[‡] |
| MDSV (6, 3) | univariée | 2.06 | 1.11 | 0.88 | 0.83 | 1.04 | 0.67 | 0.62 | 0.64 | 0.72 | 2.39 | 4.18 | 5.77 |
| | jointe | 2.13 | 1.19 | 0.91 | 0.80 | 1.06 | 0.67 | 0.65 | 0.63 | 0.64[†] | 2.34[†] | 4.14 | 5.71 |
| MDSV (10, 2) | univariée | 2.06 | 1.11 | 0.87 | 0.77 | 1.02 | 0.65 | 0.60 | 0.57 | 0.71 | 2.39 | 4.18 | 5.78 |
| | jointe | 2.10 | 1.17 | 0.87 | 0.74 | 1.05 | 0.66 | 0.61 | 0.56 | 0.64[†] | 2.35[†] | 4.14[*] | 5.71[†] |
| FTSE 100 | | | | | | | | | | | | | |
| MDSV (3, 10) | univariée | 1.06 | 0.62 | 0.43 | 0.35 | 0.58 | 0.40 | 0.31 | 0.27 | 0.29 | 1.95 | 3.61 | 5.05 |
| | jointe | 1.05 | 0.56[†] | 0.30[‡] | 0.23[‡] | 0.54[‡] | 0.34[‡] | 0.22[‡] | 0.19[‡] | 0.28 | 1.92[*] | 3.57[†] | 5.01 |
| MDSV (6, 3) | univariée | 1.06 | 0.60 | 0.40 | 0.35 | 0.58 | 0.39 | 0.30 | 0.26 | 0.30 | 1.95 | 3.60 | 5.04 |
| | jointe | 1.03[‡] | 0.55[‡] | 0.29[‡] | 0.25[†] | 0.53[‡] | 0.34[‡] | 0.22[‡] | 0.20[†] | 0.26[†] | 1.93 | 3.58 | 5.02 |
| MDSV (10, 2) | univariée | 1.05 | 0.58 | 0.37 | 0.34 | 0.57 | 0.38 | 0.28 | 0.25 | 0.29 | 1.94 | 3.59 | 5.04 |
| | jointe | 1.02[‡] | 0.54[†] | 0.29[‡] | 0.22[*] | 0.52[‡] | 0.33[‡] | 0.21[‡] | 0.18 | 0.27 | 1.92 | 3.57 | 5.03 |
| NIKKEI 225 | | | | | | | | | | | | | |
| MDSV (3, 10) | univariée | 2.10 | 1.15 | 0.80 | 0.84 | 1.06 | 0.74 | 0.58 | 0.71 | 0.81 | 2.47 | 4.15 | 5.68 |
| | jointe | 2.09 | 1.10 | 0.67[‡] | 0.53[‡] | 0.99[‡] | 0.66[‡] | 0.46[‡] | 0.48[‡] | 0.79 | 2.45 | 4.12 | 5.60[†] |
| MDSV (6, 3) | univariée | 2.12 | 1.10 | 0.73 | 0.64 | 1.05 | 0.69 | 0.53 | 0.57 | 0.81 | 2.47 | 4.16 | 5.64 |
| | jointe | 2.07[†] | 1.09 | 0.66[‡] | 0.51[†] | 0.97[‡] | 0.64[‡] | 0.43[‡] | 0.43[†] | 0.79 | 2.45 | 4.12[*] | 5.60 |
| MDSV (10, 2) | univariée | 2.11 | 1.10 | 0.72 | 0.63 | 1.03 | 0.67 | 0.49 | 0.55 | 0.80 | 2.46 | 4.15 | 5.66 |
| | jointe | 2.07[†] | 1.08 | 0.65[†] | 0.48[‡] | 0.98[‡] | 0.64[†] | 0.43[†] | 0.42[‡] | 0.79 | 2.45 | 4.11[†] | 5.58[‡] |

Les symboles *, [†] et [‡] signifient respectivement que le critère est significativement inférieur à celui de la forme univariée aux seuils de 10%, 5% et 1% en utilisant le test de Diebold-Mariano.

Ces résultats concordent avec la littérature financière. En effet, le carré des rendements financiers et les variances réalisées sont deux mesures imparfaites de la variance des prix des actifs (ou des portefeuilles d'actifs). Donc, le fait de modéliser simultanément ces deux processus permet de mieux accéder à la volatilité et donc de mieux l'estimer et la prédire.

3.7.2. Prédiction des variances réalisées

Dans cette section, nous comparons les performances des formes jointe et univariée du modèle MDSV pour prédire la variance réalisée. Le tableau 3.14 présente les trois mesures de performance des prévisions : les écarts quadratique et absolu moyens et la quasi-vraisemblance pour différents horizons. Pour chaque indice financier, chaque mesure de performance et chaque horizon, les valeurs les plus faibles sont en gras. La signification statistique des différences de critères (entre la forme jointe et la forme univariée) est évaluée par le test de Diebold-Mariano. Pour ce test, la forme univariée est considérée comme modèle de référence.

Notons que ce tableau présente les résultats ne tenant pas compte de l'effet levier. En effet, pour effectuer les prévisions de variances réalisées à un horizon supérieur à un jour dans la forme univariée, il est nécessaire de faire une hypothèse sur la dynamique des rendements financiers tel qu'expliqué dans le premier paragraphe de la section 3.5.2. Nous comparons donc les résultats n'intégrant pas l'effet levier afin de mesurer si la forme jointe présente un avantage par rapport à une forme univariée modélisant les variances réalisées. À titre indicatif, les résultats tenant compte de l'effet levier sont donnés au tableau B.4 de l'annexe B.

TABLE 3.14 – Prévision des variances réalisées : forme univariée ou forme jointe (sans effet levier)

| | | RMSFE | | | | MAFE | | | | QLIK | | | | |
|--------------------|---|-----------|-------------|-------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 | |
| S&P 500 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.47 | 0.44 | 0.43 | 0.41 | 0.21 | 0.22 | 0.31 | 0.35 | -0.37 | 1.38 | 3.27 | 4.86 |
| | | jointe | 0.48 | 0.45 | 0.44 | 0.45 | 0.22 | 0.23 | 0.32 | 0.39 | -0.37 | 1.38 | 3.26 | 4.88 |
| MDSV (6, 3) | { | univariée | 0.47 | 0.42 | 0.39 | 0.35 | 0.21 | 0.20 | 0.24 | 0.26 | -0.37 | 1.37 | 3.28 | 4.96 |
| | | jointe | 0.48 | 0.43 | 0.40 | 0.37 | 0.21 | 0.21 | 0.27 | 0.31 | -0.37 | 1.37 | 3.22 | 4.91 |
| MDSV (10, 2) | { | univariée | 0.47 | 0.42 | 0.39 | 0.33 | 0.21 | 0.20 | 0.24 | 0.25 | -0.37 | 1.37 | 3.25 | 4.88 |
| | | jointe | 0.49 | 0.43 | 0.39 | 0.35 | 0.21 | 0.21 | 0.26 | 0.28 | -0.37 | 1.37 | 3.26 | 4.89 |
| NASDAQ 100 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.59 | 0.51 | 0.47 | 0.45 | 0.27 | 0.28 | 0.33 | 0.38 | 0.03 | 1.73 | 3.53 | 5.10 |
| | | jointe | 0.60 | 0.51 | 0.45 [†] | 0.41 [‡] | 0.27 | 0.28 [†] | 0.32 [†] | 0.34 [‡] | 0.03 | 1.74 | 3.52 [†] | 5.08 [‡] |
| MDSV (6, 3) | { | univariée | 0.60 | 0.51 | 0.44 | 0.40 | 0.27 | 0.27 | 0.31 | 0.32 | 0.02 | 1.73 | 3.52 | 5.09 |
| | | jointe | 0.59 | 0.51 | 0.45 | 0.41 | 0.27 | 0.27 | 0.31 | 0.35 | 0.02 | 1.73 | 3.53 | 5.10 |
| MDSV (10, 2) | { | univariée | 0.59 | 0.51 | 0.44 | 0.38 | 0.27 | 0.26 | 0.28 | 0.28 | 0.02 | 1.73 | 3.52 | 5.10 |
| | | jointe | 0.59 | 0.51 | 0.44 | 0.39 | 0.27 | 0.26 | 0.29 | 0.31 | 0.02 | 1.73 | 3.52 | 5.09 |
| FTSE 100 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.62 | 0.34 | 0.23 | 0.20 | 0.26 | 0.18 | 0.17 | 0.16 | 0.16 | 1.78 | 3.42 | 4.85 |
| | | jointe | 0.56 | 0.35 | 0.28 | 0.32 | 0.24 [†] | 0.19 | 0.21 | 0.27 | 0.13 [†] | 1.78 | 3.44 | 4.92 |
| MDSV (6, 3) | { | univariée | 0.59 | 0.32 | 0.21 | 0.16 | 0.25 | 0.17 | 0.15 | 0.12 | 0.15 | 1.78 | 3.42 | 4.85 |
| | | jointe | 0.54 | 0.32 | 0.23 | 0.21 | 0.23 [†] | 0.18 | 0.18 | 0.18 | 0.13 | 1.77 | 3.43 | 4.86 |
| MDSV (10, 2) | { | univariée | 0.58 | 0.32 | 0.21 | 0.16 | 0.25 | 0.17 | 0.15 | 0.12 | 0.15 | 1.78 | 3.43 | 4.84 |
| | | jointe | 0.54 | 0.31 | 0.22 | 0.18 | 0.23 [‡] | 0.17 | 0.16 | 0.14 | 0.13 | 1.77 [†] | 3.43 | 4.85 |
| NIKKEI 225 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.43 | 0.33 | 0.28 | 0.25 | 0.20 | 0.18 | 0.21 | 0.23 | -0.11 | 1.54 | 3.27 | 4.75 |
| | | jointe | 0.42 | 0.33 | 0.29 | 0.30 | 0.20 | 0.18 | 0.22 | 0.26 | -0.11 | 1.54 | 3.28 | 4.78 |
| MDSV (6, 3) | { | univariée | 0.42 | 0.31 | 0.25 | 0.20 | 0.20 | 0.16 | 0.17 | 0.17 | -0.11 | 1.53 | 3.25 | 4.71 |
| | | jointe | 0.41 | 0.31 | 0.26 | 0.23 | 0.19 * | 0.17 | 0.19 | 0.21 | -0.12 | 1.54 | 3.27 | 4.75 |
| MDSV (10, 2) | { | univariée | 0.41 | 0.30 | 0.24 | 0.20 | 0.19 | 0.16 | 0.17 | 0.18 | -0.12 | 1.53 | 3.25 | 4.72 |
| | | jointe | 0.41 | 0.31 | 0.25 | 0.21 | 0.19 | 0.16 | 0.18 | 0.19 | -0.12 | 1.54 | 3.27 | 4.74 |

Les symboles *, [†] et [‡] signifient respectivement que le critère est significativement inférieur à celui de la forme univariée aux seuils de 10%, 5% et 1% en utilisant le test de Diebold-Mariano.

D'après le tableau 3.14, la forme univariée produit de meilleures prévisions que la forme jointe (100 cas sur 144). Ce résultat concorde avec le fait que la connaissance des données de rendements financiers n'améliore pas la prévision des variances réalisées contrairement au fait que la connaissance des variances réalisées améliore les prévisions des carrés de rendements financiers (tableau 3.13).

D'après le tableau 3.14, à très court terme (prévision d'un jour), dans 24 cas sur 36, la forme jointe du modèle MDSV produit de meilleures prévisions que la forme univariée. Par exemple, en dehors des spécifications ($N = 3, K = 10$) et ($N = 10, K = 2$) ajustées sur le NASDAQ 100, la forme jointe produit des QLIK à l'horizon d'un jour plus faibles que la forme univariée. De plus, le test de Diebold-Mariano indique que les MAFE à l'horizon d'un jour produits par la forme jointe des trois spécifications sont significativement (au seuil 1%) plus faibles que ceux produits par la forme univariée sur les données du FTSE 100.

3.7.3. Gestion du risque

Pour la gestion du risque, nous présentons les statistiques du test de couverture conditionnelle effectué sur les prévisions de valeur à risque de niveau 95% des formes jointe et univariée du modèle MDSV sur les quatre séries de données.

TABLE 3.15 – Gestion du risque : forme univariée ou forme jointe

| (N, K) | (3, 10) | | (6, 3) | | (10, 2) | |
|------------|---------|-------|--------|-------|---------|-------|
| | joint | univ | joint | univ | joint | univ |
| S&P 500 | 0.023* | 0.527 | 0.050 | 0.854 | 0.074 | 0.980 |
| NASDAQ 100 | 0.547 | 0.725 | 0.664 | 0.981 | 0.558 | 0.980 |
| FTSE 100 | 0.183 | 0.180 | 0.404 | 0.180 | 0.404 | 0.370 |
| NIKKEI 225 | 0.714 | 0.643 | 0.758 | 0.981 | 0.449 | 0.933 |

univ : forme univariée, joint : forme jointe. Le symbole * signifie que le critère est significativement supérieur à la statistique de test au seuil de 5%.

D'après le tableau 3.15, à l'exception de la spécification jointe avec peu de composantes ($N = 3, K = 10$) ajustée sur le S&P 500, toutes les spécifications produisent des valeurs à risque à 95% valides sur toutes les séries de données.

Les valeurs à risque conditionnelles prévues par les trois spécifications du modèle MDSV sur les données de l'indice S&P 500 (figure 3.7) montrent que la forme univariée est moins conservatrice que la forme jointe. En effet, la figure 3.7 montre que les valeurs à risque produites par les formes univariées sont plus extrêmes que celles produites par les formes jointes. De plus, selon les tests, les valeurs à risque produites par les formes univariées sont plus près du 95% quantile que celles produites par les formes jointes. Les valeurs à risques produites par les formes jointes ne sont probablement pas assez extrêmes. Elles ont donc plus que 5% des chances d'être atteintes, ce qui fait que les statistiques des tests associés sont plus élevées et que l'hypothèse nulle est plus prône à être rejetée. Cela explique aussi que le nombre de violations des formes jointes est plus grand que le nombre de violations des formes univariées.

La figure 3.7 nous indique aussi que les violations ne sont pas regroupées uniquement aux périodes de fortes volatilités (début et fin 2018) pour les deux formes (jointe et univariée) du modèle MDSV. Selon Konderla *et al.* (2017), les événements extrêmes tels que les crises financières sont difficilement capturables

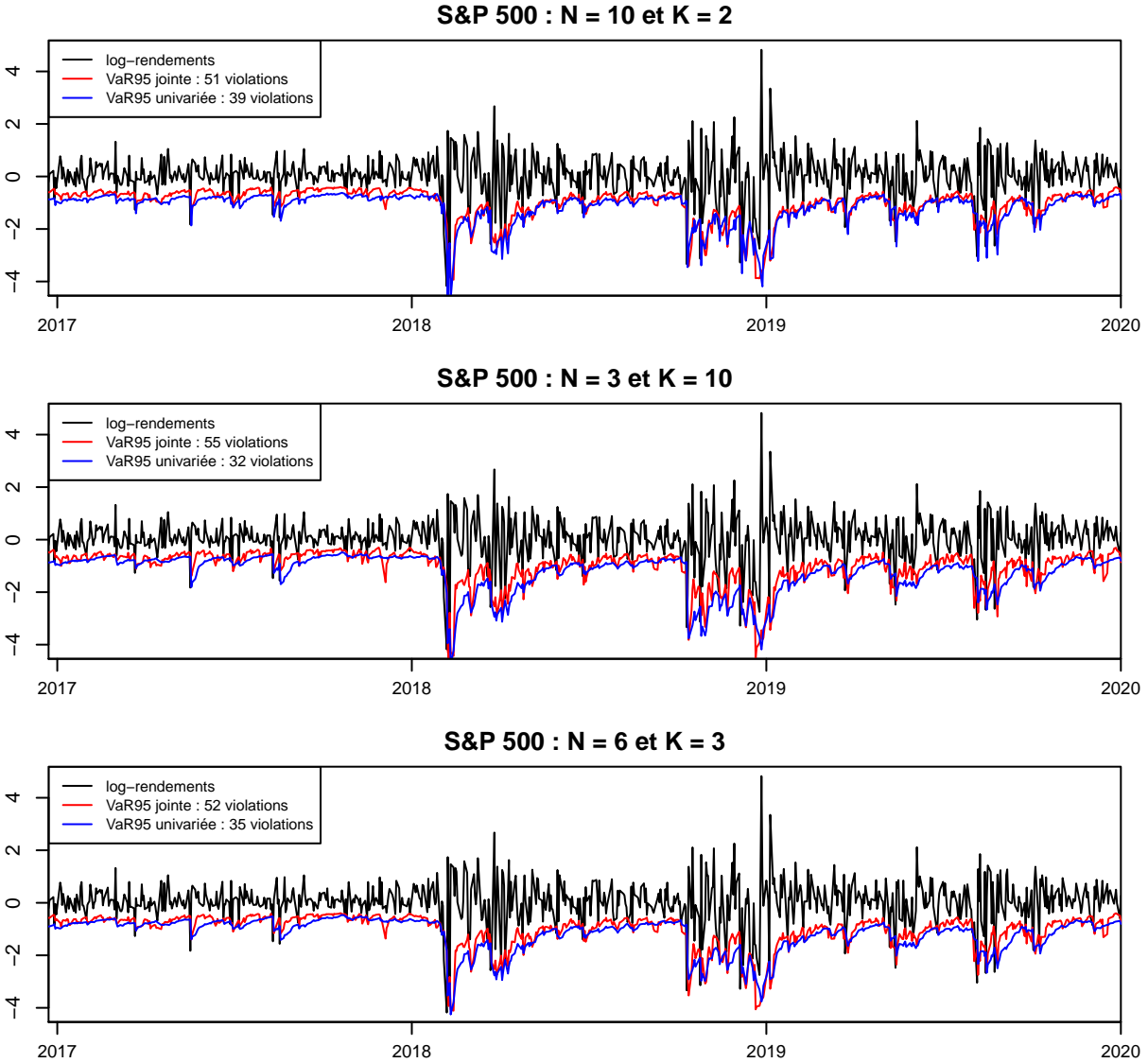


FIGURE 3.7 – Valeurs à risque conditionnelles produites sur l’indice S&P 500 par les formes jointe et univariée des trois spécifications du modèle MDSV sur la période du 22 décembre 2016 au 31 décembre 2019 (échantillon de taille 5016).

par la valeur à risque. Cependant, les valeurs à risque conditionnelles prévues par le modèle MDSV conduisent à éviter les violations pendant les moments de crises.

3.8. Analyse de l’ajustement du modèle MDSV joint sur l’indice S&P 500

Dans cette section nous analysons l’ajustement du modèle MDSV joint avec effet levier sur l’indice financier S&P 500. Plus précisément, nous étudions les paramètres estimés puis l’adéquation de la fonction d’autocorrélation théorique obtenue avec ces paramètres à celle empiriquement obtenue sur les données.

Nous étudions également l'effet levier et la distribution stationnaire de la volatilité. L'annexe B présente les résultats obtenus sur les trois autres indices financiers (NASDAQ 100, FTSE 100 et NIKKEI 225).

3.8.1. Estimation des paramètres

Le tableau 3.16 montre les estimations des paramètres par la méthode du maximum de vraisemblance de la forme jointe du modèle MDSV pour les trois spécifications $(N = 3, K = 10)$, $(N = 6, K = 3)$ et $(N = 10, K = 2)$ sur les données de l'indice financier S&P 500⁵ dans le cas où l'effet levier est pris en compte. Pour interpréter les valeurs, on doit se rappeler que chaque composante peut être vue comme un flux d'informations reçues influençant la volatilité du prix de l'actif. De plus, ces composantes ont une persistance décroissante $(\phi^{(i)})_{i=1, 2, \dots, N}$. À chaque date t , chaque composante a ainsi un impact de $(\nu_j)_{j=1, 2, \dots, K}$ qui est déterminé par l'évolution de la chaîne de Markov. Les probabilités $(\pi_j)_{j=1, 2, \dots, K}$, quant à elles, représentent les proportions de temps que chaque composante passe dans l'état ν_j . Le paramètre φ mesure l'élasticité entre la variance réalisée RV_t et la variance latente V_t . Il peut donc être interprété comme la variation (en pourcentage) de la variance réalisée RV_t induite par une variation d'un pour cent de la variance latente V_t .

Le nombre moyen de périodes qu'une chaîne de Markov reste dans un état est une mesure de persistance. Ce nombre est obtenu par la formule $(1/(1 - p_{ii}))$ où p_{ii} est l'élément i sur la diagonale de la matrice de transition de cette chaîne de Markov. La persistance de chaque composante $C_{i=1, 2, \dots, N}^{(i)}$ de la chaîne de Markov $\{V_t\}$ du modèle MDSV est donc liée au paramètre $\phi^{(i)}$. À titre de rappel, la matrice de transition de chaque composante $C^{(i)}$ est donnée par

$$\mathbf{P}^{(i)} = \phi^{(i)} \mathbf{I}_K + (1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'},$$

où \mathbf{I}_K est la matrice identité de dimension $K \times K$, $\mathbf{1}_K$ est le vecteur $K \times 1$ composé uniquement de 1, $\phi^{(i)} \in [0, 1)$ est un paramètre, et $\boldsymbol{\pi}^{(i)}$ est un vecteur paramètre de probabilités. Ainsi, plus $\phi^{(i)}$ est élevé, plus long est le temps passé par la chaîne de Markov dans chaque état. Par exemple, dans l'estimation de la spécification MDSV avec trois composantes sur les données de l'indice financier S&P 500, la première composante ($\phi^{(1)} = 0.9989$) reste dans un état environ deux ans et demi en moyenne, la deuxième composante ($\phi^{(2)} = 0.9736$) environ deux mois et la troisième composante ($\phi^{(3)} = 0.5242$) environ deux jours.

Une interprétation économique intéressante peut être associée à chacune des spécifications. En effet, les composantes sont interprétées comme de nouvelles informations qui affectent la volatilité et dont le degré de persistance est décroissant. Ainsi la première composante affecte la volatilité pendant beaucoup plus longtemps que la seconde composante et ainsi de suite. Dans le cas du modèle avec beaucoup de composantes $(N = 10, K = 2)$, l'impact de chaque composante est soit négatif (réduisant la volatilité), soit positif (augmentant la volatilité). Ainsi, sur les données S&P 500, chaque composante de la spécification avec beaucoup de composantes $(N = 10, K = 2)$ peut contribuer soit à réduire la volatilité de son quart (avec une probabilité de $\pi_1 = 78.6\%$) ou soit à l'augmenter du quart (avec une probabilité de $\pi_2 = 21.4\%$). Dans

5. Le tableau B.5 de l'annexe B.3 montre les estimations des paramètres pour les quatre indices financiers (S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225).

TABLE 3.16 – S&P 500 : Estimation par maximum de vraisemblance

| MDSV($N = 3, K = 10$) | | | | | | | | | |
|--|--------------------|----------------------|---------------------|-------------------|--------------|--------------|--------------|--------------|---------------|
| Composante constante : $\sigma^2 = 0.3156$ | | | | | | | | | |
| Composante chaîne de Markov : $\omega = 0.3151, a = 0.9989, b = 24.1159, \nu_0 = 0.8593$ | | | | | | | | | |
| $\phi^{(1)} = 0.9989, \phi^{(2)} = 0.9736, \phi^{(3)} = 0.5242$ | | | | | | | | | |
| ν_1 | ν_2 | ν_3 | ν_4 | ν_5 | ν_6 | ν_7 | ν_8 | ν_9 | ν_{10} |
| 0.8593 | 1.1407 | 1.5144 | 2.0105 | 2.6691 | 3.5435 | 4.7043 | 6.2454 | 8.2913 | 11.0074 |
| π_1 | π_2 | π_3 | π_4 | π_5 | π_6 | π_7 | π_8 | π_9 | π_{10} |
| 0.0332 | 0.1374 | 0.2527 | 0.2713 | 0.1871 | 0.0861 | 0.0264 | 0.0052 | 0.0006 | 0.0000 |
| Paramètres de variance réalisée : | | | | | | | | | |
| $\xi = -0.3848$ | $\varphi = 0.9671$ | $\delta_1 = -0.1171$ | $\delta_2 = 0.0826$ | $\gamma = 0.1381$ | | | | | |
| Paramètres d'effet levier : $l_1 = 0.7679, \theta_l = 0.8875$ | | | | | | | | | |
| MDSV($N = 6, K = 3$) | | | | | | | | | |
| Composante constante : $\sigma^2 = 0.2033$ | | | | | | | | | |
| Composante chaîne de Markov : $\omega = 0.2241, a = 0.9996, b = 5.3535, \nu_0 = 0.7686$ | | | | | | | | | |
| $\phi^{(1)} = 0.9996, \phi^{(2)} = 0.9978, \phi^{(3)} = 0.9880, \phi^{(4)} = 0.9376, \phi^{(5)} = 0.7082, \phi^{(6)} = 0.1577$ | | | | | | | | | |
| $\nu_1 = 0.7686, \nu_2 = 1.2314, \nu_3 = 1.9726$ | | | | | | | | | |
| $\pi_1 = 0.6020, \pi_2 = 0.3478, \pi_3 = 0.0502$ | | | | | | | | | |
| Paramètres de variance réalisée : | | | | | | | | | |
| $\xi = -0.3811$ | $\varphi = 0.9615$ | $\delta_1 = -0.1165$ | $\delta_2 = 0.0793$ | $\gamma = 0.1321$ | | | | | |
| Paramètres d'effet levier : $l_1 = 0.7357, \theta_l = 0.8701$ | | | | | | | | | |
| MDSV($N = 10, K = 2$) | | | | | | | | | |
| Composante constante : $\sigma^2 = 0.1809$ | | | | | | | | | |
| Composante chaîne de Markov : $\omega = 0.2140, a = 0.9998, b = 2.6887, \nu_0 = 0.7506$ | | | | | | | | | |
| $\phi^{(1)}$ | $\phi^{(2)}$ | $\phi^{(3)}$ | $\phi^{(4)}$ | $\phi^{(5)}$ | $\phi^{(6)}$ | $\phi^{(7)}$ | $\phi^{(8)}$ | $\phi^{(9)}$ | $\phi^{(10)}$ |
| 0.9998 | 0.9993 | 0.9982 | 0.9952 | 0.9870 | 0.9655 | 0.9099 | 0.7757 | 0.5052 | 0.1595 |
| $\nu_1 = 0.7506, \nu_2 = 1.2495$ | | | | | | | | | |
| $\pi_1 = 0.7861, \pi_2 = 0.2140$ | | | | | | | | | |
| Paramètres de variance réalisée : | | | | | | | | | |
| $\xi = -0.3896$ | $\varphi = 0.9634$ | $\delta_1 = -0.1156$ | $\delta_2 = 0.0870$ | $\gamma = 0.1476$ | | | | | |
| Paramètres d'effet levier : $l_1 = 0.7324, \theta_l = 0.8658$ | | | | | | | | | |

l'estimation des trois spécifications, le coefficient φ est très proche de l'unité. Cela indique que la variance latente V_t est très corrélée à la variance réalisée, ce qui concorde avec la littérature financière.

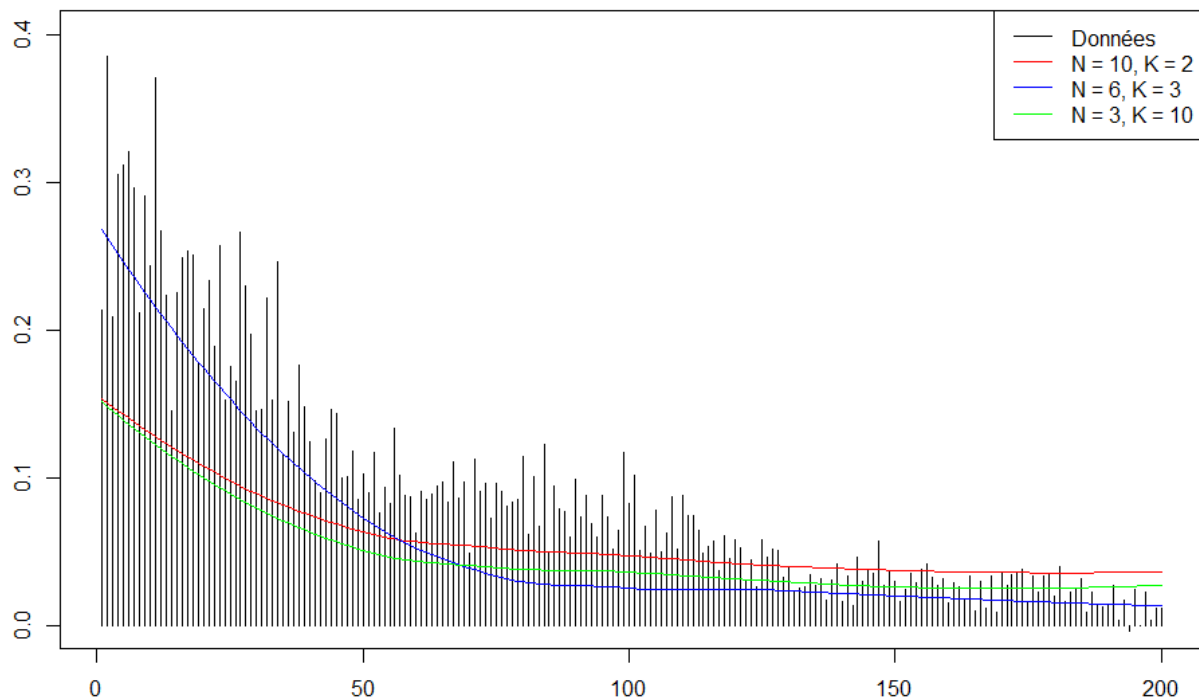


FIGURE 3.8 – Fonction d'autocorrélation empirique en noir contre les fonctions d'autocorrélation théoriques des spécifications ($N = 10, K = 2$) en rouge, ($N = 6, K = 3$) en bleu et ($N = 3, K = 10$) en vert du modèle MDSV.

3.8.2. Fonction d'autocorrélation

Une propriété importante du modèle MDSV est sa flexibilité lui permettant de reproduire des fonctions d'autocorrélation de croissance lente. Dans cette section, nous étudions l'adéquation des fonctions d'autocorrélation théorique du modèle MDSV à celle empirique. La figure 3.8 présente les fonctions d'autocorrélation théorique des trois spécifications du modèle MDSV et la fonction d'autocorrélation empirique des données de rendements financiers de l'indice S&P 500⁶. Les fonctions d'autocorrélation théorique ont été obtenues par simulation puisqu'il est difficile de calculer les moments de la composante d'effet levier $\{L_t\}$ de l'équation (2.5.1).

D'après la figure 3.8, nous observons une persistance de volatilité de longue durée pour chacune des trois spécifications. De plus, les fonctions d'autocorrélation théorique des trois spécifications et plus particulièrement celle avec un nombre moyen de composantes ($N = 6, K = 3$) suivent raisonnablement bien celle empirique.

6. La figure B.1 montre les fonctions d'autocorrélation pour les autres indices financiers.

3.8.3. Analyse de l'effet levier

L'effet levier joue un rôle important dans la modélisation des données d'indices financiers. Le modèle MDSV permet de tenir compte de cet effet levier à travers sa structure multiplicative (section 2.5). Dans cette section, nous analysons l'effet levier obtenu par l'estimation par la méthode du maximum de vraisemblance sur les données de l'indice S&P 500 dans la spécification ($N = 3, K = 10$) du modèle MDSV. Les effets levier produits par les trois spécifications du modèle MDSV sont similaires étant donné que les estimations des paramètres de la composante effet levier (l, θ) sont semblables (table 3.16). La figure 3.9 montre la valeur du coefficient l_i ($i = 1, \dots, 70$) de la composante effet levier. Cette figure peut être interprétée comme le temps nécessaire à un marché financier pour réagir complètement à un rendement négatif. Nous observons que les rendements passés, sur une fenêtre d'environ 50 jours, sont pertinents pour construire l'effet levier dans le modèle MDSV.

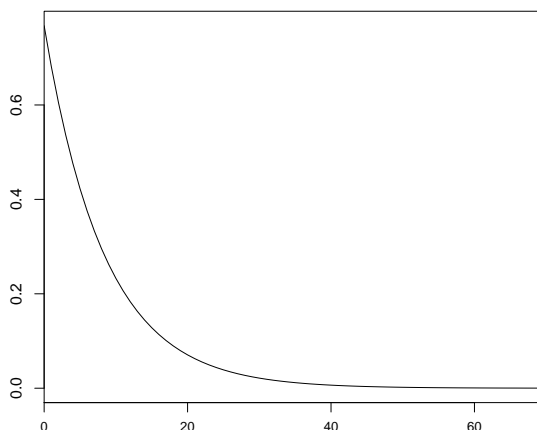


FIGURE 3.9 – Paramètre l_i .

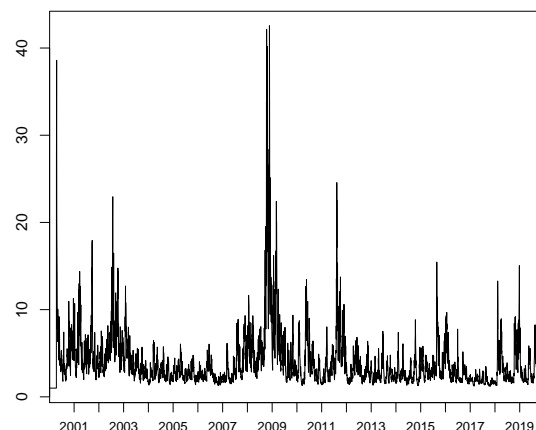


FIGURE 3.10 – L_t .

La figure 3.10, quant à elle, illustre la valeur de l'effet levier (L_t) dans le temps. D'après cette figure 3.10, les valeurs les plus élevées de l'effet levier correspondent aux dates d'événements pertinents tels que les crises financières (krach boursier en 2002, la crise des sub-primes en 2008, la crise de la dette souveraine européenne en 2011). Nous constatons que les effets de ces événements ne se résorbent qu'après plusieurs mois.

3.8.4. Distribution stationnaire de la volatilité

Il est bien connu que les rendements financiers présentent une asymétrie dans leur densité (faits stylisés, section 1.1). La figure 3.11 présente les distributions stationnaires des volatilités pour les trois spécifications du modèle MDSV.

D'après cette figure, les distributions de volatilité produites par les trois spécifications du modèle MDSV présentent une asymétrie positive. De plus, l'asymétrie semble plus prononcée à mesure que le nombre de

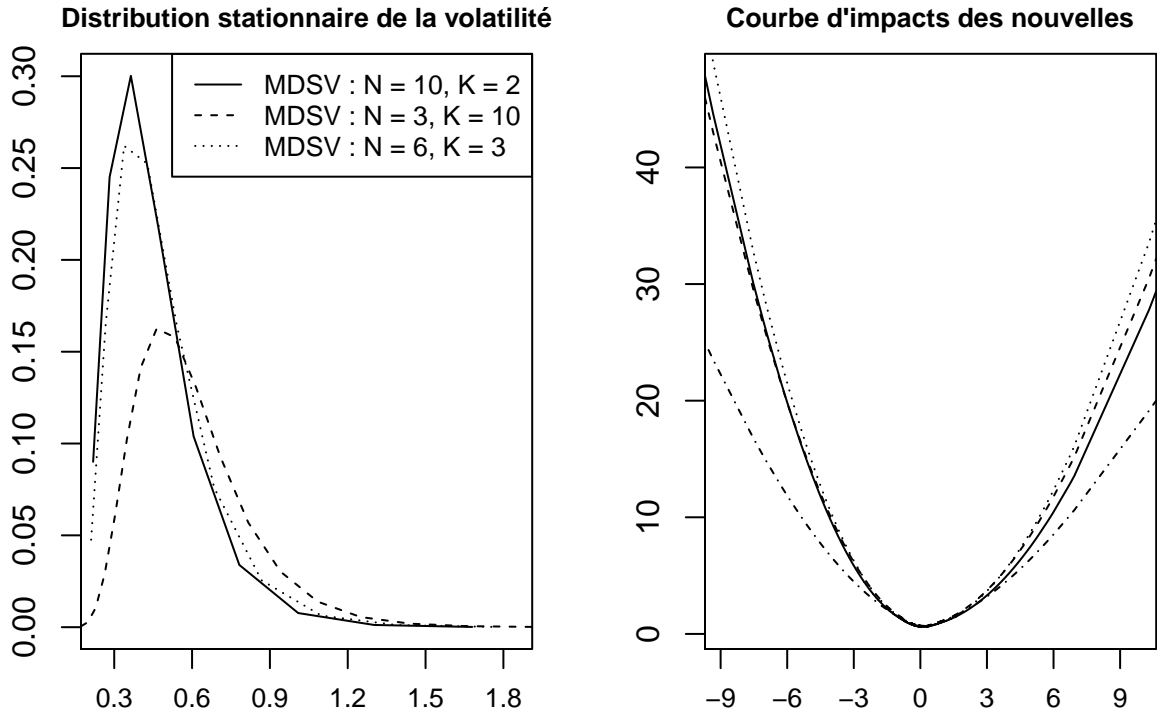


FIGURE 3.11 – Distribution stationnaire de la volatilité (à gauche) et courbe d’impacts de nouvelles (à droite). Les tracés en trait plein, en pointillés et en tirets désignent respectivement les spécifications $(N = 10, K = 2)$, $(N = 6, K = 3)$ et $(N = 3, K = 10)$ du modèle MDSV. À droite, le tracé en tirets et pointillés désigne le modèle Real EGARCH.

composantes augmente. Le modèle avec peu de composantes $(N = 3, K = 10)$ a donc plus de chance de produire des volatilités élevées que le modèle avec beaucoup de composantes $(N = 10, K = 2)$. À contrario, le modèle avec beaucoup de composantes $(N = 10, K = 2)$ a plus de chances de produire de faibles volatilités qu’un modèle avec peu de composantes $(N = 3, K = 10)$.

La figure 3.11 présente aussi la courbe d’impact des nouvelles (*news impact curve*) introduite par Engle et Ng (1993) pour mesurer comment de nouvelles informations affectent la volatilité des rendements. C’est une représentation graphique de la volatilité du lendemain (V_{t+1}) par rapport au choc des nouvelles actuelles (r_t) . La courbe d’impact des nouvelles permet de visualiser l’influence des chocs actuels sur la volatilité future. Elle reflète également l’effet levier (un choc négatif produit une volatilité plus élevée qu’un choc positif) pour les modèles GARCH asymétriques tels que les modèles EGARCH et GJR (Engle et Ng, 1993, Takahashi et al., 2013). Pour représenter la courbe d’impact des nouvelles du modèle MDSV, dans lequel la volatilité est latente, on se base sur les probabilités lissées $(\mathbb{P}(V_{t+1} | r_{1:n}, RV_{1:n}))$, où n est la taille de l’échantillon. Ces probabilités lissées sont calculées par les techniques standards de filtrage développées pour les modèles à chaîne de Markov cachée (ex., Hamilton, 1994, chapitre 22). On considère ensuite la volatilité la plus probable à chaque date afin de représenter la courbe d’impact des nouvelles. Ainsi, la figure 3.11 trace la

volatilité future (V_{t+1}) par rapport au choc du rendement actuel (r_t) de -10% à 10%. Elle montre les courbes d'impact des nouvelles du modèle Real EGARCH et des trois spécifications du modèle MDSV.

Il est bien connu que la courbe d'impact des nouvelles des modèles de type GARCH est en forme de U. Nous pouvons constater que les trois spécifications du modèle MDSV produisent des courbes d'impact des nouvelles ayant une forme de U comme celle du modèle Real EGARCH, mais présentent une plus grande volatilité pour les valeurs extrêmes des rendements financiers. Notons également que toutes les courbes d'impact des nouvelles reflètent une asymétrie négative (effet levier), tel que prévu par la littérature financière. Cette asymétrie semble plus prononcée dans le cas des modèles MDSV. Par exemple, un rendement actuel de -9% entraîne une volatilité future d'environ 22 pour le Real EGARCH et d'environ 40 pour le MDSV. Par contre, un rendement positif de même ampleur (9%) entraîne une volatilité future plus faible (environ 16 pour le Real EGARCH et environ 25 pour le MDSV). Ce qui correspond à une diminution relative de 27.3% pour le Real EGARCH et de 37.5% pour le MDSV.

3.9. Discussion

Le présent chapitre propose une analyse empirique du modèle MDSV introduit au chapitre 2. Une particularité de ce modèle est de permettre non seulement l'estimation des modèles univariés de rendements financiers ou de variances réalisées, mais aussi l'estimation conjointe de ces deux processus. À travers l'analyse empirique, nous avons exposé les performances d'ajustement, de prévision de densité et de volatilité, puis de gestion du risque du modèle MDSV sous ses formes univariées modélisant les rendements financiers et les variances réalisées, ainsi que sous sa forme jointe. En accord avec la littérature financière, le fait de tenir compte de l'effet levier permet d'améliorer grandement les performances d'ajustement du modèle MDSV, quel que soit sa forme. En terme de prévision, prendre en compte l'effet levier permet d'améliorer principalement les prévisions à court terme (faibles horizons). Les formes univariées du modèle MDSV produisent des performances d'ajustement et de prévision similaires à celles des modèles de référence. Cependant, le modèle MDSV produit de meilleures prévisions à long terme. En effet, le test de Diebold-Mariano nous permet de conclure que le modèle MDSV domine significativement le modèle FHMV pour des horizons élevés. La forme jointe du modèle MDSV montre que celui-ci domine significativement les modèles de référence, aussi bien en termes d'ajustement que de prévision de volatilité et de log-vraisemblance prédictive. Nous avons aussi montré que le modèle MDSV produit des valeurs à risque conditionnelles de niveau 95% valides. Cela prouve que ce modèle est adapté pour la gestion du risque financier.

Dans ce chapitre, nous avons aussi étudié empiriquement l'avantage comparatif d'un modèle joint par rapport à un modèle univarié. Nous avons conclu que la forme jointe du modèle MDSV est plus performante que la forme univariée, aussi bien pour l'ajustement que pour la prévision de la densité et de la volatilité. La forme jointe permet, en particulier, d'améliorer les prévisions à court terme. Cependant, pour des prévisions à long terme de variances réalisées, il n'apparaît y pas avoir un apport significatif à utiliser la forme jointe par rapport à la forme univariée. Bien que les formes jointe et univariée produisent toutes les deux des valeurs à

risque conditionnelles de niveau 95% valides, la forme jointe semble être plus conservatrice (produisant des valeurs à risque moins extrêmes) que la forme univariée.

En terme d'ajustement, la performance du modèle MDSV augmente à mesure qu'on augmente le nombre d'états des composantes jusqu'à une certaine valeur, puis se met à diminuer légèrement. On peut donc déduire qu'un nombre d'états par composante trop faible ou un trop grand entraîne une baisse de la performance d'ajustement. De même, pour un nombre d'états fixé, les performances d'ajustement du modèle MDSV augmentent à mesure qu'on augmente le nombre de composantes, puis se met à diminuer. Il faut donc envisager la détermination du nombre de composantes et/ou d'états à l'aide des données au lieu de leur attribuer une valeur *ex ante*. L'objectif du chapitre 4 est la détermination du nombre de composantes d'un modèle factoriel à chaîne de Markov en utilisant les données financières.

Chapitre 4

Le modèle iFHMV : *infinite Factorial Hidden Markov Volatility*

À travers les trois précédents chapitres, nous avons montré que les modèles à chaîne de Markov cachée de haute dimension, et plus particulièrement les modèles factoriels à chaîne de Markov cachée (FHMM), constituent un outil puissant de modélisation des données financières. Ces modèles factoriels à chaîne de Markov cachée supposent que la volatilité présente dans les données financières s’obtient par le produit d’un grand nombre (mais un nombre fini) de chaînes de Markov appelées composantes. Dans ce chapitre, au lieu de fixer *ex ante* un nombre de composantes, nous autorisons sa détermination en utilisant les données. Pour ce faire, nous employons une approche bayésienne basée sur le processus de buffet indien (IBP, *Indian Buffet Process*).

4.1. Motivation et justification

En finance, les modèles factoriels à chaîne de Markov (MSM de [Calvet et Fisher \(2004\)](#), FHMV de [Augustyniak et al. \(2019\)](#) et MDSV présenté dans les trois premiers chapitres de cette thèse) sont des modèles à volatilité stochastique à temps discret. Ces différents modèles supposent que la volatilité totale est le produit d’un grand nombre (fini) de composantes (chaînes de Markov), présentant ainsi une structure factorielle. Dans leurs analyses de simulations, [Calvet et Fisher \(2004\)](#) montrent que le fait d’avoir très peu de composantes ou trop de composantes pénalise la fonction de vraisemblance du modèle MSM. Cependant, la pénalité est plus faible lorsque le nombre de composantes est élevé. Ce même résultat est obtenu avec le modèle MDSV dans la section 2.4. C’est pour cette raison qu’en pratique, on utilise le plus grand nombre de composantes possible sous la contrainte de la charge de calcul. Par exemple, les modèles FHMV et MSM utilisent 10 composantes ayant chacune 2 états et se limitent ainsi à travailler avec un maximum de 1024 états pour représenter les volatilités. En fixant ainsi un nombre de composantes, on suppose qu’il n’existe aucune incertitude liée à ce nombre. Cependant, nous avons montré dans le chapitre 2 que la vitesse de décroissance de la fonction d’autocorrélation dans les modèles factoriels est liée au nombre de composantes. En augmentant le nombre de composantes, on a plus de flexibilité mais la fonction d’autocorrélation décroît plus rapidement. Fixer le plus grand nombre composantes possible pourrait donc s’avérer contre productif.

En se basant sur des contributions importantes (Ghahramani et Griffiths, 2006, Teh *et al.*, 2007) portant sur les processus du buffet indien (*Indian Buffet Process*), les modèles factoriels à chaîne de Markov cachée (FHMM) ont été étendus à une dimension infinie (Gael *et al.*, 2009, Valera *et al.*, 2015). Le modèle *infinite Factorial Hidden Markov Model* (iFHMM), un modèle bayésien non-paramétrique, permet une très grande flexibilité de la distribution conditionnelle, qui peut changer dans le temps. En d’autres termes, au lieu de fixer, préalablement, un nombre de composantes comme dans les modèles MSM, FHMV ou MDSV, on autorise les données à déterminer le nombre idéal de composantes nécessaires, qui pourrait varier dans le temps. Gael *et al.* (2009) et Valera *et al.* (2015) montrent que ces modèles performant mieux que leurs versions en dimension finie. De plus, l’application de ces modèles aux domaines de la reconnaissance vocale et de la désagrégation de puissance indiquent que le fait d’inférer le nombre de composantes à partir des données améliore les résultats comparativement au fait de fixer un nombre *a priori*.

Dans le domaine financier, Hartman et Heaton (2011), Dufays (2015), Maheu et Yang (2016) et Liu et Maheu (2018) ont étendu les modèles à chaîne de Markov classiques à une dimension infinie et ont observé une augmentation de performance en termes d’estimation et de log-vraisemblance prédictive. Selon Liu et Maheu (2018), cette structure classique de la chaîne de Markov définie pour modéliser la volatilité, bien qu’ayant une dimension infinie, ne permet pas de capter une forte persistance; son modèle ne peut donc pas être appliqué à des données de haute fréquence. Rappelons que la structure multiplicative des modèles factoriels à chaîne de Markov cachée permet de capter simultanément les différentes nuances de fréquences (hautes, basses et intermédiaires). Un modèle factoriel de dimension infinie pourrait donc permettre de capter cette forte persistance.

Dans ce chapitre, nous proposons une extension du modèle FHMV au cadre infini, ce qui vient compléter les modèles de Augustyniak *et al.* (2019) et Liu et Maheu (2018). Ce processus, que nous nommons iFHMV (*infinite Factorial Hidden Markov Volatility*), apporte plus de flexibilité aux deux structures proposées par Augustyniak *et al.* (2019) et Liu et Maheu (2018) en les combinant. Il autorise les données à déterminer le nombre de composantes nécessaire à leur modélisation. Étant un modèle bayésien, il nous permet d’obtenir des probabilités *a posteriori* pouvant être utilisées par un gestionnaire de portefeuille afin de prendre des décisions éclairées. Dans la section 4.4, nous utilisons le processus iFHMV pour modéliser les données de rendements financiers d’indices financiers et d’actions. Ce processus nous permet également de comparer deux approches (bayésienne et fréquentiste) de modélisation des données financières.

4.2. Présentation du processus iFHMV

Le processus iFHMV est un processus en temps discret noté $\{V_t\}$ représentant un facteur de risque à valeurs positives dans un modèle économétrique. On définit V_t de sorte à être le produit de plusieurs composantes, chacune étant une chaîne de Markov :

$$V_t = \prod_{k=1}^N C_t^{(k)}. \quad (4.2.1)$$

Cette structure factorielle est présentée plus en détails dans les sous-sections 4.2.1 et 4.2.2. Les données financières sont reliées au processus iFHMV dans la section de l'analyse empirique (section 4.4).

4.2.1. Cas constant : Nombre de composantes N fixe

Soit $\{s_t^{(1)}\}, \{s_t^{(2)}\}, \dots, \{s_t^{(N)}\}$, un ensemble de N chaînes de Markov d'espace d'états $\{0, 1\}$ et soit N nombres réels $p_i \in (0, 1)$, $i \in \{1, 2, \dots, N\}$, tels que la matrice définie par

$$\mathbf{P}_i = \begin{pmatrix} 1 - p_i & p_i \\ p_i & 1 - p_i \end{pmatrix}$$

soit la matrice de transition de la chaîne $\{s_t^{(i)}\}$.

Définition 4.2.1. *On dira que la chaîne de Markov $\{s_t^{(i)}\}$ est active à la date t lorsque $s_t^{(i)} = 1$. Elle sera dite inactive dans le cas contraire, c'est-à-dire $s_t^{(i)} = 0$.*

On suppose qu'à la date $t = 0$, toutes les chaînes sont inactives. Donc $s_0^{(i)} = 0$ pour tout $i = 1, 2, \dots, N$. On définit chaque composante $C_t^{(i)}$ de V_t dans l'équation (4.2.1) de sorte qu'elle prenne la valeur $c_i > 0$ lorsque la chaîne de Markov $\{s_t^{(i)}\}$ est active et la valeur 1¹ lorsqu'elle ne l'est pas :

$$C_t^{(i)} = c_i^{s_t^{(i)}} = \begin{cases} 1 & \text{si } s_t^{(i)} = 0 \\ c_i & \text{si } s_t^{(i)} = 1 \end{cases} .$$

Ainsi, la volatilité V_t vérifie la relation

$$V_t = \prod_{i=1}^N C_t^{(i)} = \prod_{i=1}^N c_i^{s_t^{(i)}} .$$

Remarque 4.2.1. *Le cas constant du modèle iFHMV partage des similarités avec les autres modèles factoriels présentés dans cette thèse. Par exemple, en dimension finie (N fixe),*

- *le modèle iFHMV est un cas particulier du modèle MDSV présenté aux chapitres 2 et 3. En effet, en fixant $K = 2$, $\phi^{(i)} = 1 - 2p_i$, $\boldsymbol{\pi}^{(i)} = \left(\frac{1}{2} \quad \frac{1}{2}\right)'$ et $\boldsymbol{\nu}^{(i)} = \left(1 \quad c_i\right)'$ dans le modèle MDSV, on obtient le modèle iFHMV en dimension finie.*
- *le modèle FHMV de Augustyniak et al. (2019) est un cas particulier du modèle iFHMV. En effet, en supposant, dans le modèle iFHMV en dimension finie, que les probabilités d'activation p_i sont constantes et égales entre elles puis que les volatilités c_i sont liées par la relation $c_i = (1 - \theta_c) + \theta_c c_{i-1}$, $i = 1, 2, \dots, N$ où $\theta_c \in (0, 1)$, on se ramène à la composante persistante du modèle FHMV.*

Remarque 4.2.2.

L'intérêt est particulièrement porté sur les chaînes de Markov $\{s_t^{(i)}\}$. Désignons ainsi par \mathcal{S} la matrice binaire de dimension $T \times N$ contenant les états de toutes les chaînes de Markov $\{s_t^{(i)}\}_{t=1, \dots, T}$ à chaque

1. Le choix de la valeur 1 est fait pour faciliter l'estimation et l'identification des paramètres du modèle. Ce choix est en accord avec les données financières que nous étudions qui ont une variance échantillonnale sensiblement proche de l'unité.

temps t :

$$\mathcal{S} = \left(s_t^{(i)} \right)_{\substack{t=1,\dots,T \\ i=1,\dots,N}} = \begin{pmatrix} s_1^{(1)} & s_1^{(2)} & \dots & s_1^{(N)} \\ s_2^{(1)} & s_2^{(2)} & \dots & s_2^{(N)} \\ \vdots & \vdots & \vdots & \vdots \\ s_T^{(1)} & s_T^{(2)} & \dots & s_T^{(N)} \end{pmatrix}.$$

Pour chacune des chaînes $\{s_t^{(i)}\}_{t=1,\dots,T}$, dénotons le nombre de transitions de types $0 \rightarrow 0$, $0 \rightarrow 1$, $1 \rightarrow 0$ et $1 \rightarrow 1$ successivement par $n_{00}^{(i)}$, $n_{01}^{(i)}$, $n_{10}^{(i)}$, $n_{11}^{(i)}$.

Proposition 4.2.1 (Loi de \mathcal{S}).

(i) Conditionnellement aux probabilités p_1, \dots, p_N , la loi de \mathcal{S} est donnée par :

$$\mathbb{P}(\mathcal{S} \mid p_1, \dots, p_N) = \prod_{i=1}^N p_i^{n_{01}^{(i)} + n_{10}^{(i)}} (1 - p_i)^{n_{00}^{(i)} + n_{11}^{(i)}},$$

où $\mathbb{P}(\cdot)$ représente la loi.

(ii) Si on suppose que les probabilités p_i suivent une distribution bêta de paramètres $\frac{\alpha}{N}$ et 1 a priori, avec $\alpha > 0$, alors la loi marginale de \mathcal{S} est donnée par :

$$\mathbb{P}(\mathcal{S} \mid \alpha) = \prod_{i=1}^N \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}, \quad (4.2.2)$$

où $\Gamma(\cdot)$ est la fonction Gamma et $n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} = T$ pour tout $i = 1, \dots, N$.

DÉMONSTRATION. Voir l'annexe C.1. □

Remarque 4.2.3. Une permutation des composantes $\{C_t^{(1)}\}, \{C_t^{(2)}\}, \dots, \{C_t^{(N)}\}$, correspondant à une permutation des colonnes de \mathcal{S} , ne modifie pas les distributions conditionnelles et marginales de \mathcal{S} et ne modifie pas la valeur de la volatilité V_t à la date t . On suppose ici et dans la suite du chapitre qu'une permutation sur les colonnes de \mathcal{S} correspond aussi à une permutation des composantes $\{C_t^{(1)}\}, \{C_t^{(2)}\}, \dots, \{C_t^{(N)}\}$.

La remarque 4.2.3 indique que le modèle factoriel tel que défini par l'équation (4.2.1) est échangeable dans les colonnes. En d'autres termes, l'ordre des chaînes de Markov $\{s_t^{(i)}\}$ n'influence pas le modèle. Nous définissons alors la classe des matrices \mathcal{S} qui ne modifient pas le modèle.

Définition 4.2.2. Soit une matrice binaire \mathcal{S} ayant T lignes. On associe à chaque colonne $\left(s_t^{(i)}\right)$ de \mathcal{S} , son encodage binaire défini par le nombre $2^{T-1}s_1^{(i)} + 2^{T-2}s_2^{(i)} + \dots + s_T^{(i)}$. La forme lof (left-ordered form) de \mathcal{S} est la matrice ayant les mêmes colonnes que \mathcal{S} , mais ordonnées de façon décroissante selon leurs encodages.

On dira aussi que \mathcal{S} est une matrice lof si elle est égale à sa forme lof.

Le graphique 4.1 présente un exemple de matrice binaire, accompagnée de sa forme lof.

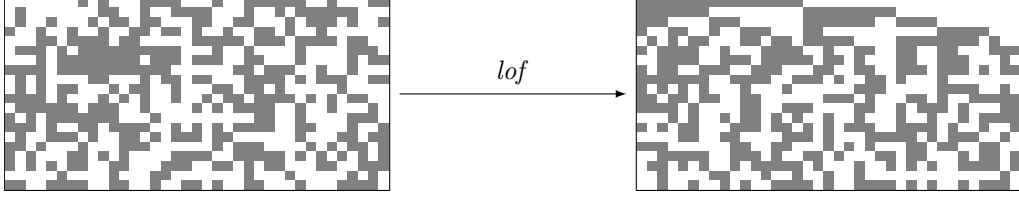


FIGURE 4.1 – Forme *lof* d’une matrice binaire ayant 20 lignes (désignant les dates) et 37 colonnes (désignant les chaînes de Markov). Les cases grises indiquent que la chaîne de Markov est active.

4.2.2. Nombre infini de composantes : N tendant vers l’infini

L’objectif dans cette section est de faire tendre le nombre de composantes N de la matrice \mathcal{S} vers l’infini. Dans l’équation (4.2.2), la probabilité d’une telle matrice est bien sûr nulle lorsque N tend vers l’infini. Cependant, ceci n’est pas un problème puisque ce qui nous intéresse n’est pas la loi de cette unique matrice mais bien celle de toute une classe de matrices : la classe des matrices qui peuvent s’obtenir par permutation des colonnes d’une autre matrice de la même classe. En d’autres termes, nous nous intéressons à la loi de la classe des matrices qui ont la même forme *lof* lorsque le nombre de colonnes tend vers l’infini. Cette classe de matrices a aussi été utilisée par [Gael et al. \(2009\)](#) et [Valera et al. \(2015\)](#).

Supposons que \mathcal{S} soit une matrice *lof* ; notons que toute matrice peut être transformée en matrice *lof* moyennant quelques permutations des colonnes selon l’encodage binaire présenté précédemment. Nous désignons par $[\mathcal{S}]$ l’ensemble de toutes les matrices qui peuvent être ré-exprimées sous la forme \mathcal{S} en utilisant uniquement des permutations de colonnes ; on appelle $[\mathcal{S}]$ la classe d’équivalence *lof*.

Théorème 4.2.1 (Loi de $[\mathcal{S}]$). *Soit \mathcal{S} , une matrice binaire de dimension $T \times N$ dans laquelle chaque colonne d’indice i correspond aux états d’une chaîne de Markov $\{s_t^{(i)}\}$ qui prend ses valeurs dans $\{0, 1\}$ aux dates $t = 1, \dots, T$. On suppose que \mathcal{S} est une matrice *lof* ; sinon, elle peut être transformée en matrice *lof* par une permutation des colonnes selon l’encodage binaire. Désignons aussi par N_h le nombre de colonnes de \mathcal{S} ayant le même encodage binaire h et par $[\mathcal{S}]$ la classe d’équivalence *lof* associée à \mathcal{S} . On a les résultats suivants :*

(i) *Le nombre d’éléments de la classe d’équivalence $[\mathcal{S}]$ est $\frac{N!}{\prod_{h=0}^{2^T-1} N_h!}$.*

(ii) *La loi de la classe d’équivalence $[\mathcal{S}]$ est*

$$\mathbb{P}([\mathcal{S}] | \alpha) = \frac{N!}{\prod_{h=0}^{2^T-1} N_h!} \prod_{i=1}^N \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}. \quad (4.2.3)$$

(iii) *Lorsque $N \rightarrow +\infty$, la loi de la classe d’équivalence $[\mathcal{S}]$ est*

$$\lim_{N \rightarrow +\infty} \mathbb{P}([\mathcal{S}] | \alpha) = \frac{\alpha^{N_+}}{\prod_{h=0}^{2^T-1} N_h!} \exp\{-\alpha H_T\} \prod_{i=1}^{N_+} \frac{\left(n_{01}^{(i)} + n_{10}^{(i)} - 1\right)! \left(n_{00}^{(i)} + n_{11}^{(i)}\right)!}{\left(n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)}\right)!}, \quad (4.2.4)$$

où $H_T = \sum_{j=1}^T \frac{1}{j}$ et N_+ est le nombre de chaînes qui se sont activées au moins une fois entre $t = 1$ et $t = T$.

(iv) La dimension effective du modèle (N_+) est distribuée selon une loi de Poisson de paramètre αH_T .

DÉMONSTRATION. Voir annexe C.2. □

Ce modèle est analogue au processus du buffet indien (*Indian Buffet Process* de [Ghahramani et Griffiths, 2006](#) et [Teh et al., 2007](#)). Ces processus partagent une interprétation très intéressante. Pour présenter le processus du buffet indien, supposons que T clients entrent, en ordre, dans un restaurant indien avec un buffet de taille infinie sur lequel sont placés des repas en ligne. Le premier client entre dans le restaurant et se sert en prenant une portion de chaque repas en commençant à gauche du buffet et en s'arrêtant après un nombre de loi de $\text{Poisson}(\alpha)$ repas parce que son assiette devient surchargée. Un serveur reste à côté du buffet et note le nombre de personnes ayant mangé chaque repas. Le t -ième client entre dans le restaurant et commence à gauche du buffet. Au plat i , il vérifie si le client le précédant ($(t-1)$ -ième client) s'est servi ce repas.

- Si oui, il demande au serveur le nombre de personnes qui se sont servi le repas i quand celui qui les précédait s'était servi ce repas (le serveur lui répond $n_{11}^{(i)}$) et le nombre de personnes qui ne s'étaient pas servi le repas i bien que celui qui les précédait se soit servi ce repas (le serveur lui répond $n_{10}^{(i)}$). Le client se sert alors ce repas i avec probabilité $\frac{n_{11}^{(i)}}{n_{11}^{(i)} + n_{10}^{(i)}}$.
- Si non, il demande au serveur combien de personnes se sont servi ce repas i bien que celui qui les précédait ne se soit pas servi ce repas (le serveur répond $n_{01}^{(i)}$) et combien de personnes ne se sont pas servi ce repas i lorsque celui qui venait avant eux ne s'était pas servi ce repas non plus (le serveur lui répond $n_{00}^{(i)}$). Le client se sert alors le repas i avec probabilité $\frac{n_{01}^{(i)}}{n_{01}^{(i)} + n_{00}^{(i)}}$.

Le client se déplace ensuite vers le repas suivant et fait exactement pareil. Après avoir passé tous les repas consommés par d'autres personnes avant lui, il essaie un nombre de $\text{Poisson}(\alpha/t)$ nouveaux repas. Si nous désignons par $N_1^{(t)}$ le nombre de nouveaux plats essayés par le t -ième client, la probabilité qu'une matrice particulière soit produite par ce processus, appelé le processus du buffet indien (*Indian Buffet Process*), est (selon [Ghahramani et Griffiths, 2006](#))

$$\frac{\alpha^{N_+}}{\prod_{t=0}^T N_1^{(t)}!} \exp\{-\alpha H_T\} \prod_{i=1}^{N_+} \frac{(n_{01}^{(i)} + n_{10}^{(i)} - 1)! (n_{00}^{(i)} + n_{11}^{(i)})!}{(n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)})!}. \quad (4.2.5)$$

Nous pouvons retrouver l'équation (4.2.4) en additionnant toutes les matrices possibles qui peuvent être générées en utilisant ce processus et qui sont dans la même classe d'équivalence *lof*; il y a exactement $\frac{\prod_{t=0}^T N_1^{(t)}!}{\prod_{h=0}^{2^T-1} N_h!}$ matrices dans la même classe d'équivalence *lof*. En multipliant cela par l'équation (4.2.5), on retrouve l'équation (4.2.4). Cette construction montre que la dimension effective du modèle (N_+) suit une loi de $\text{Poisson}(\alpha H_T)$.

4.2.3. Construction rupture de bâton (*stick-breaking*)

Bien que la construction présentée à la section 4.2.2 soit très élégante théoriquement, elle n'aide pas vraiment dans l'inférence. Nous décrivons, dans cette section, une représentation alternative du modèle qui impose un ordre aux composantes de la matrice \mathcal{S} , lié à leurs probabilités d'activation. Cette représentation, appelée « construction rupture de bâton », est fortement liée à la construction rupture de bâton (*stick-breaking*) du processus de Dirichlet. Afin de produire cette construction rupture de bâton, il faut disposer d'une série décroissante de probabilités $p_{(1)} > p_{(2)} > \dots$ dont la distribution est connue. Teh *et al.* (2007) montrent que pour un ensemble de probabilités $p_i, i = 1, 2, \dots$, distribuées indépendamment selon une loi bêta de paramètres $\frac{\alpha}{N}$ et 1 avec $\alpha > 0$, on a :

$$p_{(1)} \mid \alpha \sim \text{Beta}(\alpha, 1),$$

$$\mathbb{P}(p_{(i)} \mid p_{(1:i-1)}, \alpha) = \alpha p_{(i)}^{\alpha-1} p_{(i-1)}^{-\alpha} \mathbf{1}_{\{0 \leq p_{(i)} \leq p_{(i-1)}\}},$$

où $p_{(i)}$ est la i -ième plus grande probabilité et $p_{(1:i-1)} = \{p_{(1)}, \dots, p_{(i-1)}\}$. Mieux, les probabilités $q_{(i)}$ définies pour chaque i par

$$q_{(i)} = \begin{cases} p_{(1)} & \text{si } i = 1, \\ p_{(i)}/p_{(i-1)} & \text{si } i > 1, \end{cases}$$

sont indépendantes et identiquement distribuées selon une loi bêta de paramètres α et 1. Disposant de ces probabilités, la construction rupture de bâton peut être comprise métaphoriquement comme suit. On part d'un bâton de longueur 1. A l'itération $i = 1, 2, \dots$, on casse un morceau en un point $q_{(i)}$ par rapport à la longueur courante du bâton. Autrement dit, $q_{(i)}$ représente la proportion du bâton précédent que l'on conserve. Nous enregistrons la longueur $p_{(i)}$ du bâton que nous venons de rompre, et itérons sur ce nouveau morceau, en jetant l'autre morceau de bâton. La figure 4.2 présente une construction rupture de bâton. Il s'agira donc de tronquer la construction rupture de bâton après un certain nombre de points de ruptures et d'effectuer une inférence dans un espace réduit.

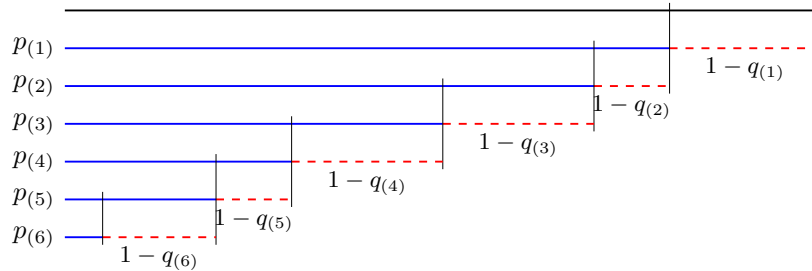


FIGURE 4.2 – Construction rupture de bâton. La ligne horizontale noire est de taille 1. A chaque itération, la ligne verticale noire représente le point de rupture, la ligne bleue représente la longueur du bâton, la ligne rouge en pointillée représente le morceau du bâton qui a été jeté. L'expression sous la ligne rouge représente la proportion du bâton qui a été jetée.

Concrètement, dans le cas de notre modèle, rappelons que p_i est la probabilité que la chaîne $\{s_t^{(i)}\}$ s'active au moins une fois à partir de l'état 0. En ordonnant ainsi les probabilités, on ordonne les chaînes de Markov dans l'ordre décroissant de leurs probabilités d'activation. L'idée est alors de supposer qu'après un certain nombre de composantes, les composantes restantes ne s'activeront pas puisque leurs probabilités d'activation sont trop faibles. Soit N_+ le niveau de troncature. On posera $p_i = 0$ pour tout $i > N_+$. En faisant cette supposition, l'influence de \mathcal{S} dans la volatilité (équation (4.2.1)) est la même que celle de la matrice réduite aux N_+ premières colonnes de \mathcal{S} étant donné que les colonnes restantes ne s'activeront pas. Ainsi, par hypothèse, si le niveau de troncature est connu, le problème se ramène à un cas fini. L'estimation de ce niveau de troncature est effectué par la méthode du *slice sampling* présenté à la section suivante.

4.3. Estimation du modèle

Dans cette section, nous présentons la procédure d'estimation du modèle iFHMV. Le modèle tel que présenté est un modèle bayésien non-paramétrique. Son estimation peut se faire par l'algorithme d'échantillonnage de Gibbs. Dans un premier temps, nous présentons concrètement la méthode d'estimation de la dimension effective du modèle (N_+). Nous présentons ensuite la procédure d'estimation du modèle par l'algorithme de Gibbs par blocs (*Block Gibbs Sampling*).

4.3.1. *Slice sampling* : Estimation de N_+

Teh *et al.* (2007) ont proposé un moyen d'appliquer le *stick-breaking* dans le cadre du processus du buffet indien. En effet, la construction rupture de bâton prédéfinit un niveau de troncature (N_+). La méthode *slice sampling* nous permet d'estimer ce niveau de troncature. Elle consiste en l'introduction d'une variable auxiliaire permettant de transformer un problème d'échantillonnage d'une distribution arbitraire en un problème d'échantillonnage d'une loi uniforme.

Soit un nombre infini de chaînes de Markov $\{s_t^{(i)}\}_{i=1,2,\dots}$ caractérisées par la série décroissante de probabilités $p_{1:\infty} = (p_{(i)})_{i=1,2,\dots}$ telles que $p_{(1)} > p_{(2)} > \dots$. On introduit une variable auxiliaire μ :

$$\mu \mid \mathcal{S}, p_{1:\infty} \sim \text{Uniforme} [0, p^*], \quad (4.3.1)$$

où p^* est une fonction de $p_{1:\infty}$ et de \mathcal{S} . On suppose que p^* est la taille du bâton associé à la dernière composante qui s'est activée au moins une fois :

$$p^* = \min_{i:\exists t, s_t^{(i)}=1} p_{(i)}.$$

La distribution jointe de \mathcal{S} , $p_{1:\infty}$ et de la variable auxiliaire μ vérifie la relation :

$$\mathbb{P}(\mu, p_{1:\infty}, \mathcal{S}) = \mathbb{P}(\mathcal{S}, p_{1:\infty}) \mathbb{P}(\mu \mid \mathcal{S}, p_{1:\infty}) = \mathbb{P}(\mathcal{S}, p_{1:\infty}) \left[\frac{1}{p^*} \mathbf{1}_{\{0 \leq \mu \leq p^*\}} \right].$$

Clairement, l'intégration de μ préserve la distribution d'origine sur $p_{1:\infty}$ et \mathcal{S} tandis que conditionnellement à $p_{1:\infty}$ et \mathcal{S} , μ est simplement échantillonné à l'aide d'une distribution uniforme. Conditionnellement à μ , aux

paramètres $c_{1:\infty} = (c_i)_{i=1,2,\dots}$ et aux données $r_{1:T} = (r_1, \dots, r_T)$ du modèle, la distribution de \mathcal{S} vérifie :

$$\mathbb{P}(\mathcal{S} \mid r_{1:T}, c_{1:\infty}, p_{1:\infty}, \mu) \propto \mathbb{P}(\mathcal{S} \mid r_{1:T}, c_{1:\infty}, p_{1:\infty}) \left[\frac{1}{p^*} \mathbf{1}_{\{0 \leq \mu \leq p^*\}} \right]. \quad (4.3.2)$$

Puisque p^* est la plus petite probabilité pour laquelle la composante est active, la condition $\mu \in [0, p^*]$ dans l'équation (4.3.2) implique que toutes les composantes dont la probabilité d'activation est plus petite que μ n'auront pas d'impact sur \mathcal{S} . Nous conviendrons alors que les colonnes de \mathcal{S} pour lesquelles les probabilités d'activation $p_{(i)}$ sont plus petite que μ sont nulles. Soit N_+ , le nombre de composantes pour lesquelles $p_{(i)} > \mu$. Cette procédure impose donc que $s_t^{(i)} = 0$ pour tout $i > N_+$ et pour tout t , nous permettant ainsi de mettre uniquement à jour les N_+ premières composantes dans l'estimation du modèle. Notons que N_+ sert de niveau de troncature dans la mesure où il limite les coûts de calcul à un nombre fini sans approximation. Finalement, la détermination du niveau de troncature (N_+) se réduit à un échantillonnage selon une loi uniforme (échantillonnage de μ).

4.3.2. Algorithme d'estimation du modèle iFHMV

Avant de présenter l'algorithme d'estimation du modèle iFHMV, nous allons relier le processus iFHMV aux données de rendements financiers. De manière similaire aux modèles financiers populaires, le rendement financier r_t à la date t est distribué selon une loi normale de moyenne 0 et de variance V_t :

$$r_t \sim \text{Normal}(0, V_t), \quad (4.3.3)$$

où le processus $V = \{V_t\}$ désigne le processus iFHMV présenté à l'équation (4.2.1).

L'algorithme d'estimation du modèle iFHMV passe par l'introduction d'une variable N^* plus grande que le niveau de troncature N_+ . Plus spécifiquement, désignons par N^* un indice tel que toutes les composantes actives aient un indice $i < N^*$ (notons que N^* lui-même est l'indice d'une composante inactive mais qu'il n'est pas nécessairement égal à $N_+ + 1$). La représentation informatique du *slice sampling* (composée des variables à mettre à jour à chaque itération) est : $\langle \mu, N^*, N_+, \mathcal{S}_{1:T,1:N^*}, p_{(1:N^*)}, c_{1:N_+}, \alpha \rangle$, où $\mathcal{S}_{1:T,1:N^*}$ désigne la sous-matrice de \mathcal{S} constituée des T premières lignes et des N^* premières colonnes, $p_{(1:N^*)} = (p_{(1)}, \dots, p_{(N^*)})$ et $c_{1:N_+} = (c_1, \dots, c_{N_+})$. Rappelons que les N_+ premières colonnes de \mathcal{S} sont les seules colonnes non nulles et donc les colonnes d'indices $N_+ + 1$ à N^* contiennent uniquement des zéros. Afin de permettre un mélange plus rapide, nous utilisons une estimation par l'échantillonnage de Gibbs par blocs (*Block Gibbs Sampling*). L'échantillonnage de Gibbs par blocs, utilisant le *slice sampling* à l'intérieur d'un bloc pour estimer N_+ , est présenté à l'algorithme 2.

L'algorithme 2 est caractérisé d'échantillonnage par blocs puisqu'à sa ligne 4, les colonnes de la matrice \mathcal{S} sont mises à jour simultanément. Cette procédure permet une convergence plus rapide de l'algorithme. Présentons maintenant plus en détails cet algorithme, qui se base sur les distributions *a priori* suivantes

Algorithme 2 Algorithme d'échantillonnage de Gibbs par blocs

- 1: **Initialisation** de N^* et de N_+ . ($N^* > N_+$).
 - 2: **Initialisation** aléatoire de $\mathcal{S}_{1:T, 1:N_+}, p_{(1:N^*)}, c_{1:(N_+)}$ et α . Les colonnes d'indices $N_+ + 1$ à N^* de \mathcal{S} contiennent uniquement des zéros.
 - 3: **loop**
 - 4: **mise à jour** de $p_{(1:(N^*-1))} = (p_{(1)}, p_{(2)}, \dots, p_{(N^*-1)})$ et de $p_{(N^*)}$. Pour $p_{(N^*)}$, en plus de prendre en compte le fait que la composante d'indice N^* est inactive (à toutes les dates), il faut également prendre en compte le fait que toutes les colonnes de \mathcal{S} au-delà de N^* sont également inactives (à toutes les dates).
 - 5: **mise à jour** de μ . Si la valeur de μ est telle que $N_+ \geq N^*$, alors il faut ajouter des composantes inactives (à toutes les dates) à \mathcal{S} jusqu'à ce que $N_+ < N^*$
 - 6: **mise à jour** de $\mathcal{S} = \left(s_t^{(i)} \right)_{\substack{i=1, 2, \dots, N^* \\ t=1, 2, \dots, T}}$
 - 7: **mise à jour** de $c_{1:N_+} = (c_1, c_2, \dots, c_{N_+})$
 - 8: **mise à jour** de α
 - 9: **fin loop**.
-

pour tout $i = 1, 2, \dots$,

$$c_i \sim \text{Inv-Gamma}(a_c, b_c)^2,$$

$$\alpha \sim \text{Gamma}(a_\alpha, b_\alpha).$$

Initialisations : Il s'agit d'initialiser les variables N^* , N_+ , $\mathcal{S}, p_{(1:N^*)}, c_{1:(N_+)}$ et α . Il est à noter que N^* est strictement supérieur à N_+ et que seules les N_+ premières colonnes de \mathcal{S} ne sont pas nulles. À titre indicatif, pour nos analyses empiriques, nous initialisons $N_+ = 1, N^* = 2, p_{(1:N^*)} = (0.5, 0), c_1 = 1$ et \mathcal{S} est une matrice colonne dont la longueur est la taille de la série de données modélisées et les valeurs sont aléatoirement échantillonnées dans $\{0, 1\}$.

Mise à jour de $p_{(1:(N^*-1))}$ et $p_{(N^*)}$: Les distributions *a posteriori* des $(N^* - 1)$ probabilités ordonnées de façon décroissante sont données par

$$p_{(i)} \mid \mathcal{S}, p_{(i-1)}, p_{(i+1)} \sim \text{Beta} \left(n_{01}^{(i)} + n_{10}^{(i)}, n_{00}^{(i)} + n_{11}^{(i)} + 1 \right) \mathbf{1}_{\{p_{(i+1)} < p_{(i)} < p_{(i-1)}\}},$$

où $p_{(0)} = 1$. La dérivation de cette distribution *a posteriori* est présentée à l'annexe C.4. Pour la probabilité $p_{(N^*)}$, en plus de prendre en compte le fait que la composante d'indice N^* est inactive, nous devons également prendre en compte le fait que toutes les colonnes de \mathcal{S} au-delà de N^* sont également inactives. La proposition 4.3.1 suivante donne la distribution *a posteriori* de $p_{(N^*)}$.

Proposition 4.3.1 (Loi *a posteriori* de $p_{(N^*)}$). *La distribution a posteriori de la probabilité $p_{(N^*)}$ de la composante N^* pour laquelle les indices de toutes les composantes actives i vérifient $i < N^*$, est donnée par*

$$\mathbb{P} \left(p_{(N^*)} \mid p_{(N^*-1)}, \mathcal{S}_{:, \geq N^*} = 0 \right) \propto \exp \left(\alpha \sum_{t=1}^T \frac{1}{t} (1 - p_{(N^*)})^t \right) p_{(N^*)}^{\alpha-1} (1 - p_{(N^*)})^T \mathbf{1}_{\{0 \leq p_{(N^*)} \leq p_{(N^*-1)}\}}, \quad (4.3.4)$$

où $\mathcal{S}_{:, \geq N^*}$ représente la matrice composée de toutes les lignes de \mathcal{S} et de toutes les colonnes d'indice supérieur ou égal à N^* ; c est une matrice de colonnes inactives.

DÉMONSTRATION. Voir annexe C.3. □

Mise à jour de μ : Comme décrit à l'équation (4.3.1), μ est échantillonné par une loi uniforme

$$\mu \mid \mathcal{S}, p_{1:\infty} \sim \text{Uniforme} [0, p^*],$$

où $p^* = \min_{i:\exists t, s_t^{(i)}=1} p_{(i)}$. La mise à jour de μ est équivalente à une mise à jour de N_+ puisque N_+ est le nombre de composantes pour lesquelles $p_{(i)} > \mu$. Ainsi, à cette étape, si $p_{(N^*)} > \mu$, c'est à dire si N_+ est supérieur ou égal au nombre de colonnes N^* de \mathcal{S} , on met à jour N^* en générant autant de probabilités qu'il faut, par la distribution en (4.3.4), pour que $p_{(N^*)} < \mu$. Dans cette situation, il faut ajouter $N_+ - N^* + 1$ composantes inactives (à toutes les dates) à \mathcal{S} et ajouter des composantes au vecteur de volatilités $c_{1:N_+}$ par leur distribution *a priori*.

Mise à jour de \mathcal{S} : Chacune des colonnes de \mathcal{S} est mise à jour à travers l'algorithme *forward-backward* aussi appelé l'algorithme *Forward Filtering Backward Sampling* (FFBS) (Chib, 1996, Robert *et al.*, 1999, Scott, 2002). Cet algorithme est relié au filtre avant-arrière (*forward backward filter*) de Baum *et al.* (1970). Pour présenter cet algorithme, supposons qu'on veuille simuler la colonne d'indice i ($s_{1:T}^{(i)}$) et désignons par $\mathbf{x}_{1:T}$, le T-uplet (x_1, x_2, \dots, x_T) . Posons également $\theta = (c_{1:N_+}, p_{(1:N_+)})$ et désignons par \mathcal{S}_{-i} la matrice \mathcal{S} sans la colonne d'indice i . Le filtre commence par les étapes de filtrage avant suivantes :

Pour $t \in \{2, \dots, T\}$, supposons qu'on connaisse $\mathbb{P}(s_{t-1}^{(i)} \mid r_{1:t-1}, \mathcal{S}_{-i}, \theta)$. Par la loi des probabilités totales, nous avons :

$$\mathbb{P}(s_t^{(i)} \mid r_{1:t-1}, \mathcal{S}_{-i}, \theta) = \sum_{j=0}^1 \mathbb{P}(s_t^{(i)} \mid s_{t-1}^{(i)} = j, p_{(i)}) \times \mathbb{P}(s_{t-1}^{(i)} = j \mid r_{1:t-1}, \mathcal{S}_{-i}, \theta),$$

puisque $\mathbb{P}(s_t^{(i)} \mid r_{1:t-1}, s_{t-1}^{(i)} = j, \mathcal{S}_{-i}, \theta) = \mathbb{P}(s_t^{(i)} \mid s_{t-1}^{(i)} = j, p_{(i)})$. Ensuite,

$$\mathbb{P}(s_t^{(i)} \mid r_{1:t}, \mathcal{S}_{-i}, \theta) \propto \mathbb{P}(s_t^{(i)} \mid r_{1:t-1}, \mathcal{S}_{-i}, \theta) \times \mathbb{P}(r_t \mid \theta, s_t^{(i)}, \mathcal{S}_{-i}),$$

où $\mathbb{P}(r_t \mid \theta, s_t^{(i)}, \mathcal{S}_{-i})$ représente la distribution de la donnée r_t en fonction du vecteur de paramètres θ , de la matrice \mathcal{S}_{-i} et de l'état $s_t^{(i)}$ de la composante $C_t^{(i)}$ à la date t . Lorsque r_t est distribuée selon l'équation (4.3.3), on a :

$$\mathbb{P}(r_t \mid \theta, s_t^{(i)}, \mathcal{S}_{-i}) = \frac{1}{\sqrt{2\pi \prod_{i=1}^{N_+} c_i^{s_t^{(i)}}}} \exp\left(-\frac{r_t^2}{2 \prod_{i=1}^{N_+} c_i^{s_t^{(i)}}}\right).$$

Comme $s_0^{(i)} = 0$ pour tout i , alors pour $t = 1$,

$$\mathbb{P}(s_1^{(i)} \mid r_0, \mathcal{S}_{-i}, \theta) = \mathbb{P}(s_1^{(i)} \mid s_0^{(i)} = 0, p_{(i)}) = \begin{cases} 1 - p_{(i)} & \text{si } s_1^{(i)} = 0 \\ p_{(i)} & \text{si } s_1^{(i)} = 1 \end{cases}.$$

Désignons par $F^{(i)}$ la matrice $T \times 2$ contenant les probabilités filtrées $\mathbb{P}(s_t^{(i)} \mid r_{1:t}, \mathcal{S}_{-i}, \theta)$. Conditionnellement à \mathcal{S} , θ , $r_{1:T}$ et la $(t-1)$ -ième ligne $F_{t-1}^{(i)}$ de $F^{(i)}$, nous pouvons obtenir la t -ième ligne $F_t^{(i)}$ de $F^{(i)}$ par la relation

$$F_t^{(i)} = (F_{t-1}^{(i)} P_i) \odot d_t,$$

où d_t est le vecteur ligne contenant $\mathbb{P}(r_t \mid \theta, s_t^{(i)}, \mathcal{S}_{-i})$ et \odot est le produit de Hadamar (élément par élément).

Le filtre permet ensuite d'échantillonner les états $s_t^{(i)}$, $t = 1, \dots, T$ par l'arrière à l'aide des étapes suivantes :

- (1) Simulation de $s_T^{(i)}$ à partir de $F_T^{(i)}$,
- (2) Pour $t = T - 1, \dots, 1$, simuler $s_t^{(i)}$ à partir de $P_{s_t^{(i)} s_{t+1}^{(i)}} \times F_t^{(i)}$.

Mise à jour de $c_{1:N_+}$: Les volatilités $c_{1:N_+}$ sont échantillonnées selon leurs distributions *a posteriori* qui dépendent de la relation qui relie les données $\{r_t\}$ au processus iFHMV. Ainsi, si r_t est distribuée selon l'équation (4.3.3), les distributions *a posteriori* des volatilités $c_i, i = 1, \dots, N_+$ sont données par

$$c_i \mid \mathcal{S}, \mathbf{x}_{1:T} \sim \text{Inv-Gamma} \left(a_c + \frac{s^{(i)}}{2}, b_c + \frac{1}{2} \sum_{t=1}^T \frac{r_t^2 s_t^{(i)}}{\prod_{\substack{j=1 \\ j \neq i}}^{N_+} c_j^{s_t^{(j)}}} \right),$$

où $s^{(i)} = \sum_{t=1}^T s_t^{(i)}$. La dérivation de cette distribution *a posteriori* est présentée à l'annexe C.4.

Mise à jour de α : La distribution *a posteriori* de α est donnée par

$$\alpha \mid p_{(1:N_+)} \sim \text{Gamma} (a_\alpha + N_+ + 2, b_\alpha - \log p_{(N_++1)}) .$$

La dérivation de cette distribution *a posteriori* est présentée à l'annexe C.4.

4.4. Analyse empirique du modèle

Cette section concerne l'analyse empirique du modèle iFHMV. Il s'agit d'estimer le modèle iFHMV sur des données d'indices financiers et d'actions. Les données utilisées pour les estimations sont les séries journalières de rendements des indices financiers NASDAQ 100 et FTSE 100 disponibles sur le site Web de *Oxford-Man Institute Quantitative Finance* et présentées à la section 3.1 du chapitre 3. Ces données s'étendent sur la période du 1^{er} janvier 2000 au 31 décembre 2019. Puisque le modèle iFHMV n'intègre pas l'effet levier, nous utilisons également des données de rendements financiers d'actions s'étendant sur la période du 1^{er} janvier 2003 au 31 décembre 2012. Il s'agit des actions GES (*General Electric Stock*) et CS (*Citigroup Stock*). En effet, contrairement aux données d'indices financiers qui présentent toujours un effet levier, les données d'actions en présentent très peu. La figure 4.3, qui présente différentes fonctions d'autocorrélation empiriques concernant les données NASDAQ 100, FTSE 100, GES et CS, permet d'étudier la présence d'un effet levier. Dans ces graphiques, on dira qu'il y a une présence d'effet levier lorsque les quantités $\text{Corr}[|r_{t+h}|, \max(-r_t, 0)], h > 0$ (représentées en noir) sont supérieures aux quantités $\text{Corr}[|r_{t+h}|, \max(r_t, 0)], h > 0$ (représentées en rouge). L'analyse de la figure 4.3 indique que les données d'indices financiers (NASDAQ 100 et FTSE 100) présentent un effet levier tandis que les données d'actions (GES et CS) en présentent très peu. Nous considérons ainsi les données des deux indices NASDAQ 100 et FTSE 100 et les données des deux actions GES et CS pour mesurer la performance du modèle iFHMV sur ces catégories de données. Le tableau 4.1 rapporte les statistiques descriptives des séries journalières de rendements financiers (r_t) pour les deux indices financiers et les deux actions.

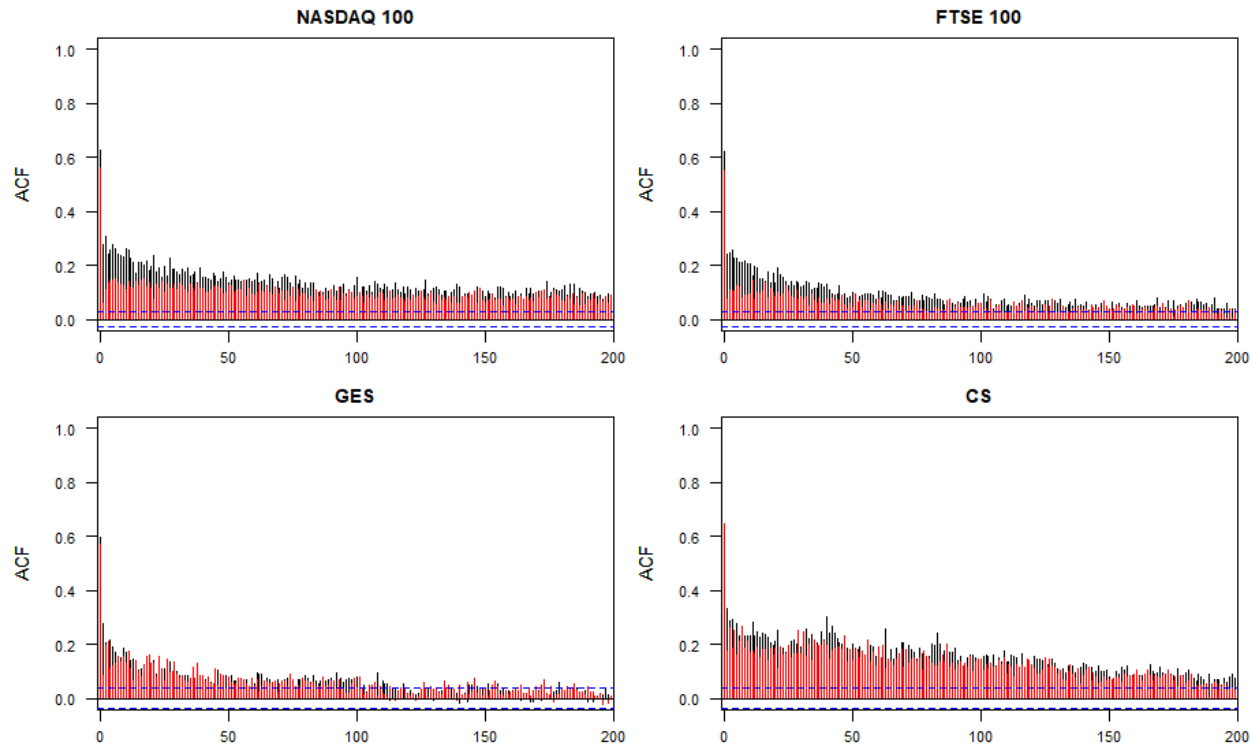


FIGURE 4.3 – Fonctions d'autocorrélation empiriques $\text{Corr}[|r_{t+h}|, \max(-r_t, 0)]$ en noir, $\text{Corr}[|r_{t+h}|, \max(r_t, 0)]$ en rouge. La première ligne présente les ACF pour les indices NASDAQ 100 et FTSE 100 et la deuxième ligne présente les ACF pour les actions GES et CS.

TABLE 4.1 – Statistiques descriptives

| Données | n | Moy. | SD | Asym. | Apla. | Min. | Q1 | Med. | Q3 | Max. |
|------------|------|------|------|-------|-------|--------|-------|------|------|-------|
| NASDAQ 100 | 5012 | 0.00 | 1.55 | -0.05 | 9.09 | -10.24 | -0.64 | 0.07 | 0.72 | 13.26 |
| FTSE 100 | 5042 | 0.00 | 1.13 | -0.17 | 9.53 | -8.93 | -0.53 | 0.03 | 0.57 | 9.48 |
| GES | 2510 | 0.00 | 1.65 | -0.35 | 13.01 | -11.15 | -0.67 | 0.05 | 0.76 | 12.41 |
| CS | 2510 | 0.00 | 3.14 | -2.30 | 30.98 | -36.35 | -0.82 | 0.17 | 1.06 | 19.49 |

Statistiques descriptives des indices NASDAQ 100 et FTSE 100 sur la période du 1^{er} janvier 2000 au 31 décembre 2019 et des actions GES et CS sur la période du 1^{er} janvier 2003 au 31 décembre 2012. n est la taille d'échantillon, Min. est le minimum de la série, Q1 est le premier quartile, Moy. est la moyenne, Med. est la médiane, SD est l'écart-type, Asym. mesure l'asymétrie, Apla. mesure l'aplatissement, Q3 est le troisième quartile et Max. est le maximum.

Les rendements financiers des quatre données financières ont une queue épaisse et une asymétrie négative. Cependant, les données d'actions (GES et CS) présentent une asymétrie plus prononcée et sont plus aplaties que les données d'indices (NASDAQ 100 et FTSE 100). Notons également que les rendements de l'action de Citigroup sont plus volatils.

4.4.1. Estimation du modèle (*in-sample*)

Le rendement financier r_t est relié au processus iFHMV de volatilité V_t à la date t de façon similaire à d'autres modèles financiers de la littérature par la relation :

$$r_t = \sqrt{V_t} \epsilon_t,$$

où $\{\epsilon_t\}$ est un processus d'innovation indépendant et identiquement distribué (*i.i.d*) selon une loi normale de moyenne nulle et de variance 1 et $\{V_t\}$ est le processus iFHMV présenté à la section 4.2.

L'estimation du modèle iFHMV sur les rendements financiers des données d'indices (NASDAQ 100 et FTSE 100) et d'actions (GES et CS) est effectuée par l'échantillonneur de Gibbs par blocs présenté dans la section 4.3. Pour cette estimation, 15 000 itérations sont effectuées et les statistiques sont présentées sur les 3 750 dernières itérations. La figure 4.4 présente les diagrammes à bandes des valeurs de N_+ inférées pour les quatre séries financières.

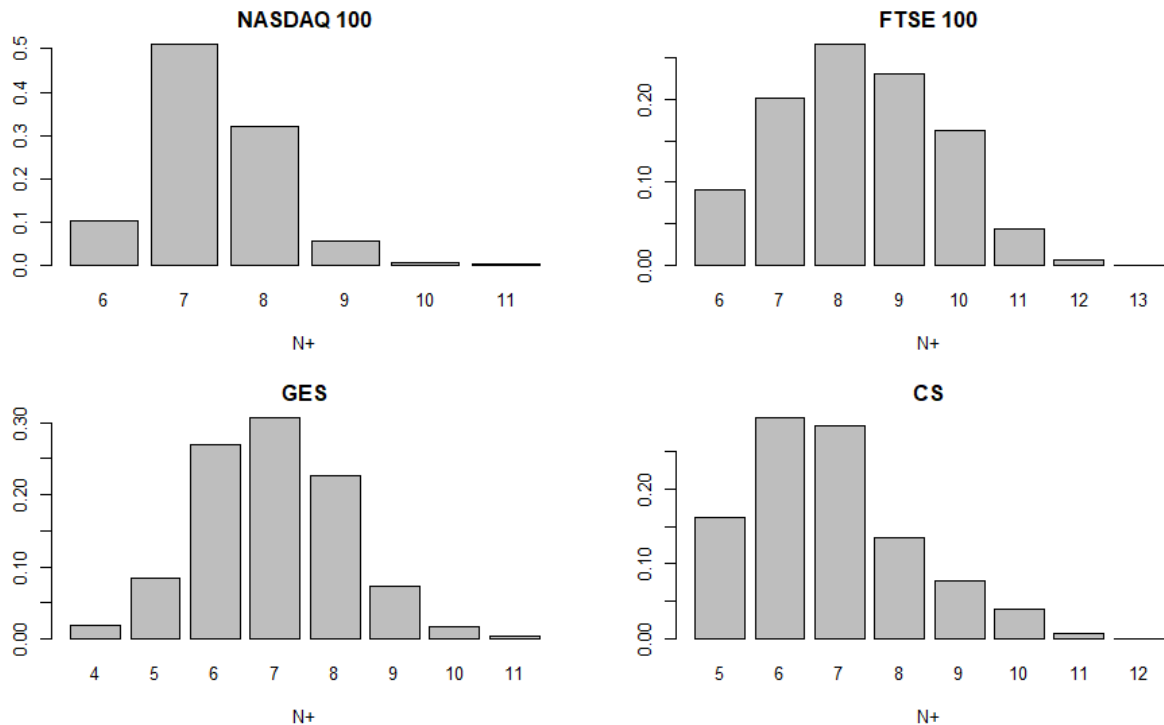


FIGURE 4.4 — Nombre de composantes N_+ inféré par la méthode de l'échantillonnage de Gibbs par blocs en utilisant l'algorithme 2. La première ligne présente les diagrammes à bandes pour les indices NASDAQ 100 et FTSE 100 et la deuxième ligne présente les diagrammes à bandes pour les actions GES et CS. Des informations complémentaires sur les résultats d'estimation sont fournies à l'annexe C.5.

Ces diagrammes à bandes indiquent qu'il existe une incertitude sur le nombre de composantes N dans les données des quatre séries financières. Notons aussi que pour les deux indices NASDAQ 100 et FTSE 100, le nombre de composantes ayant la plus grande probabilité *a posteriori* est respectivement de 7 et 8.

Pour les deux actions GES et CS, ce nombre est respectivement 7 et 6. Cela confirme que le fait de fixer N *a priori* comme dans le cas des modèles FHMV de [Augustyniak et al. \(2019\)](#), MSM de [Calvet et Fisher \(2004\)](#) et CDRS de [Fleming et Kirby \(2013\)](#) ($N = 10$ dans ces modèles), pourrait entraîner une perte de performance. Le modèle iFHMV ne prend pas une valeur *ex ante* pour le nombre de composantes, il intègre automatiquement l'incertitude présente dans le nombre de composantes pour les estimations et les prévisions. Remarquons aussi que le nombre de composantes inféré pour le NASDAQ 100 est concentré autour du nombre 7 (plus de 50% du temps). Cela indique que le fait de prendre $N = 7$ permettrait probablement de faire une bonne estimation dans le cas où l'on voudrait utiliser une valeur fixe pour le nombre de composantes.

Afin de mesurer les performances d'estimation du modèle iFHMV, nous le comparons à un autre modèle bayésien à chaîne de Markov cachée : le modèle iHMM (*infinite Hidden Markov Model*) proposé par [Maheu et Yang \(2016\)](#). Ces auteurs supposent que la moyenne et la variance des rendements financiers sont liées à une chaîne de Markov de dimension infinie modélisée par un processus de Dirichlet hiérarchique :

$$\begin{cases} \Gamma \mid \eta & \sim \text{DP}(\eta, H), \quad \theta_j = (\mu_j, \sigma_j^2) \sim H, \quad j = 1, 2, \dots, \\ P_j \mid \alpha, \Gamma & \sim \text{DP}(\alpha, \Gamma), \quad j = 1, 2, \dots, \\ r_t \mid s_t, \theta & \sim \text{Normal}(\mu_{s_t}, \sigma_{s_t}^2), \\ s_t \mid s_{t-1}, P & \sim P_{t-1}, \end{cases}$$

où $\{s_t\}$ représente la chaîne de Markov de matrice de transition P dont les lignes $P_j, j = 1, 2, \dots$ vérifient $\mathbb{E}[P_j] = \Gamma$ ³. La notation DP désigne le processus de Dirichlet et H est une distribution *a priori* vérifiant $H(\mu) \equiv \text{Normal}(m, v^2)$ et $H(\sigma^2) \equiv \text{Inv-Gamma}(v_0, s_0)$. Les paramètres α et η sont appelés paramètres de concentration et contrôlent le nombre d'états actifs du modèle. De grandes valeurs de α et η favorisent un grand nombre d'états et de petites valeurs produisent des modèles plus parcimonieux.

Les modèles iFHMV et iHMM se ressemblent dans le sens qu'ils sont deux modèles à chaîne de Markov cachée prenant en compte le fait qu'il existe une incertitude quant au nombre d'états de la chaîne de Markov. Cependant, le modèle iHMM suppose qu'il existe une seule chaîne de Markov ayant un nombre infini d'états tandis que le modèle iFHMV suppose qu'il existe une infinité de chaînes de Markov ayant chacune deux états. Dans cette section, nous comparons les performances d'estimation de ces deux modèles. Pour cela, nous effectuons 15 000 itérations pour estimer les deux modèles (iFHMV et iHMM) puis nous comparons les statistiques des 3 750 dernières itérations. Le graphique 4.5 présente les estimations par noyau (*kernel density*) des log-vraisemblances produites à l'aide des 3 750 dernières itérations des modèles iFHMV et iHMM.

La figure 4.5 indique que les densités de log-vraisemblances inférées par les modèles iFHMV et iHMM sont très différentes en ce qui concerne les données d'indices financiers et sont semblables en ce qui concerne les données d'actions. Sur les données d'indices, le modèle iHMM produit des log-vraisemblances plus grandes et

3. Pour que ce soit plus concret, prenons un exemple simple. Considérons une chaîne de Markov à trois états et supposons $\Gamma = (0.15, 0.57, 0.28)$. Chaque ligne P_j de la matrice de transition est obtenue par $P_j \sim \text{Dirichlet}(\Gamma)$, où $\text{Dirichlet}(\Gamma)$ est la distribution de Dirichlet de paramètre Γ et $\mathbb{E}[P_j] = \Gamma$. Ainsi, par exemple, les lignes échantillonnées peuvent être $P_1 = (0.34, 0.64, 0.02)$, $P_2 = (0.51, 0.26, 0.23)$ et $P_3 = (0.01, 0.50, 0.49)$.

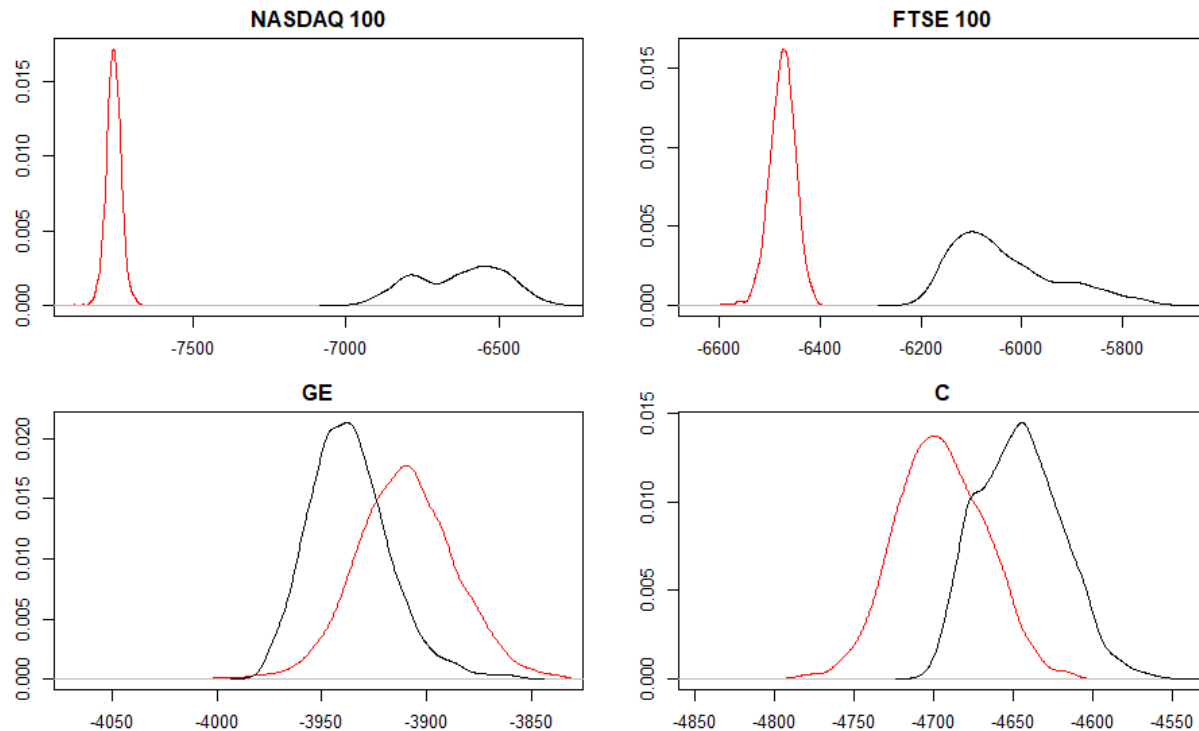


FIGURE 4.5 – Densités des log-vraisemblances calculées à l’aide des 3 750 dernières itérations de l’estimation des modèles iFHMV (rouge) et iHMM (noir). La première ligne présente les densités des log-vraisemblances pour les indices NASDAQ 100 et FTSE 100 et la deuxième ligne présente les densités des log-vraisemblances pour les actions GES et CS.

plus dispersées que le modèle iFHMV. Par contre, sur les données d’actions, les densités de log-vraisemblances inférées par les deux modèles s’entrecroisent et présentent une volatilité similaire. Le tableau 4.2 présente les statistiques descriptives des log-vraisemblances produites à l’aide des 3 750 dernières itérations des modèles iFHMV et iHMM.

TABLE 4.2 – Comparaison d’estimation

| Séries | n | Q1 : 25% | | Moy. | | SD | | Q3 : 75% | |
|------------|------|----------------|----------------|----------------|----------------|--------|-------|----------------|----------------|
| | | iHMM | iFHMV | iHMM | iFHMV | iHMM | iFHMV | iHMM | iFHMV |
| NASDAQ 100 | 5012 | -6754.7 | -7773.4 | -6627.1 | -7758.0 | 145.00 | 24.40 | -6511.5 | -7742.2 |
| FTSE 100 | 5012 | -6113.0 | -6491.0 | -6033.3 | -6474.9 | 104.51 | 25.66 | -5972.1 | -6457.5 |
| GES | 2510 | -3950.0 | -3925.6 | -3936.7 | -3910.2 | 19.37 | 23.89 | -3925.3 | -3894.8 |
| CS | 2510 | -4666.1 | -4714.9 | -4645.8 | -4695.6 | 26.41 | 28.48 | -4627.7 | -4675.6 |

n est la taille d’échantillon, Q1 est le premier quartile, Moy. est la moyenne, SD est l’écart-type et Q3 est le troisième quartile. Les valeurs les plus élevées sont en gras.

Le tableau 4.2 confirme l’analyse du graphique 4.5. Remarquons que la variation des log-vraisemblances produites par le modèle iFHMV est similaire peu importe le type de données (indices financiers ou actions). Par contre, le modèle iHMM produit des log-vraisemblances très dispersées pour les indices financiers et plutôt regroupées pour les données d’actions. Une raison qui pourrait expliquer pourquoi les log-vraisemblances

produites par le modèle iHMM sont plus élevées que celles produites par le modèle iFHMV est que le modèle iHMM ne suppose pas que les moyennes dans les différents états sont nulles, contrairement au modèle iFHMV. Le modèle iHMM, étant plus flexible que le modèle iFHMV, ce n'est pas surprenant que les résultats du modèle iHMM sur les données d'estimation (*in-sample*) soient meilleurs que ceux du modèle iFHMV. Dans la section 4.4.2, nous évaluons les performances de prévision (*out-of-sample*) de ces modèles afin de déterminer si cette flexibilité se traduit par une meilleure performance du modèle.

4.4.2. Prévision du modèle (*out-of-sample*)

En se basant sur les données de rendements financiers, nous étudions le pouvoir de prédiction du modèle iFHMV. Pour ce faire, nous considérons les trois dernières années de la base de données, correspondant à 756 jours boursiers, comme échantillon de prévision. Une estimation glissante du modèle iFHMV est effectuée en considérant toute l'information disponible jusqu'à la date initiale d'estimation. Le modèle est cependant ré-estimé à chaque trois mois (63 jours boursiers). Ainsi, à chaque date de l'échantillon de prévision, s'il s'agit d'une date de ré-estimation, le modèle iFHMV est ré-estimé en faisant 1 000 itérations de l'échantillonneur de Gibbs par blocs dans lequel on initialise les variables de l'estimation $(\mu, N^*, N_+, \mathbf{S}_{1:T,1:N^*}, p_{(1:N^*)}, c_{1:N_+}, \alpha)$ par leurs valeurs obtenues à la période de ré-estimation précédente. Aussi, à chaque date de l'échantillon de prévision, en considérant les paramètres estimés à la dernière date d'estimation comme valeurs initiales, $M = 100$ itérations de l'échantillonneur de Gibbs par blocs sont effectuées. Ces M itérations sont utilisées pour effectuer les prévisions. Les quatre mesures de performance pour la prévision des modèles que nous utilisons sont : la log-vraisemblance prédictive, l'écart quadratique moyen de prévision (RMSFE, *Root Mean Squared Forecast Error*), l'écart absolu moyen de prévision (MAFE, *Mean Absolute Forecast Error*) et la quasi-vraisemblance (QLIK, *Quasi-Likelihood*).

Afin de mesurer la performance de prévision du modèle iFHMV, nous le comparons au modèle iHMM de Maheu et Yang (2016) et à son compétiteur naturel : le modèle FHMV avec 10 composantes tel que présenté à la section 1.3. Étant donné que les modèles iFHMV et iHMM ne prennent pas en compte l'effet levier, nous considérons aussi la version du modèle FHMV qui n'intègre pas l'effet levier. Le modèle FHMV adopte une approche fréquentiste d'estimation et la procédure de calcul des quatre mesures est présentée à la section 3.2.1. En ce qui concerne le modèle iHMM, la procédure de calcul des quatre mesures est similaire à celle du modèle iFHMV.

4.4.2.1. La log-vraisemblance prédictive

La log-vraisemblance prédictive est la log-densité des futurs rendements financiers étant donné le modèle et l'information disponible à la date d'évaluation. Dans cette approche, qui a entre autres été considérée par Maheu et McCurdy (2011), Maheu et Yang (2016) et Liu et Maheu (2018), il est question de calculer la log-vraisemblance du rendement du lendemain (r_{t+1}) conditionnellement à l'information disponible actuellement.

Dans une approche bayésienne comme la nôtre, la quantité ($\ell_t = \log \mathbb{P}(r_{t+1} \mid r_{1:t}, c_{1:\infty}, p_{1:\infty})$) est estimée par :

$$\begin{aligned} \ell_t &\approx \sum_{m=1}^M \log \left[\mathbb{P} \left(r_{t+1}, \mathcal{S}_{t+1,1:N_+^{(m)}} \mid r_{1:t}, c_{1:N_+^{(m)}}, p_{1:N_+^{(m)}} \right) \right], \\ &= \sum_{m=1}^M \log \left[\mathbb{P} \left(r_{t+1} \mid \mathcal{S}_{t+1,1:N_+^{(m)}}, r_{1:t}, c_{1:N_+^{(m)}}, p_{1:N_+^{(m)}} \right) \mathbb{P} \left(\mathcal{S}_{t+1,1:N_+^{(m)}} \mid r_{1:t}, c_{1:N_+^{(m)}}, p_{1:N_+^{(m)}} \right) \right], \end{aligned}$$

où M désigne le nombre d'itérations effectuées, $N_+^{(m)}$ représente le nombre de composantes inféré à l'itération m et $\mathcal{S}_{t+1,1:N_+^{(m)}} = \left(s_{t+1}^{(i)} \right)_{i=1, \dots, N_+^{(m)}}$ représente la $(t+1)$ -ième ligne de la matrice binaire contenant les états des $N_+^{(m)}$ premières chaînes de Markov. Le calcul de la log-vraisemblance prédictive ℓ_t devient facile en remarquant qu'à chaque itération m , $N_+^{(m)}$ est une constante et donc le modèle devient un modèle à chaîne de Markov cachée classique. Le calcul devient similaire à celui de l'équation (1.3.1). Le résultat, $\ell_t = \log p(r_{t+1} \mid \mathcal{F}_t, \theta)$, est utilisé pour définir la log-vraisemblance prédictive sur un horizon de T jours, représentée par

$$d_T = \sum_{t=1}^T \ell_t = \sum_{t=1}^T \log \mathbb{P}(r_{t+1} \mid r_{1:t}, c_{1:\infty}, p_{1:\infty}).$$

Ici, $t = 0$ correspond à la date finale de l'échantillon d'estimation (*in-sample*), $t = T$ représente la date finale de l'échantillon de prévision (*out-of-sample*). Notons qu'une log-vraisemblance prédictive élevée correspond à une meilleure performance de prévision.

Le tableau 4.3 présente la log-vraisemblance prédictive d_T pour le modèle iFHMV et pour les modèles de référence (FHMV et iHMM). Pour chaque indice, les valeurs les plus élevées (désignant le meilleur modèle) sont en gras.

TABLE 4.3 – Log-vraisemblance prédictive

| Modèles | NASDAQ 100 | FTSE 100 | GES | CS |
|---------|---------------|---------------|----------------|----------------|
| FHMV | -958.3 | -801.0 | -1236.6 | -1580.5 |
| iHMM | -1305.8 | -929.7 | -1334.2 | -1707.0 |
| iFHMV | -962.1 | -800.4 | -1234.1 | -1578.5 |

Les valeurs les plus élevées sont en gras.

Le tableau 4.3 montre que le modèle iFHMV présente des log-vraisemblances prédictives plus élevées que celles des modèles FHMV et iHMM pour les deux actions (GES et CS) et pour l'indice FTSE 100. En ce qui concerne l'indice NASDAQ 100, le modèle FHMV présente une plus grande log-vraisemblance prédictive. Notons cependant que les log-vraisemblances prédictives produites par le modèle iFHMV sont très proches de celles produites par le modèle FHMV.

4.4.2.2. Les autres mesures de performance de prévision

L'écart quadratique moyen de prévision (RMSFE), l'écart absolu moyen de prévision (MAFE) et la quasi-vraisemblance (QLIK) nous permettent également de mesurer les performances de prévision des modèles. Ils

nous permettent ainsi de comparer le pouvoir prédictif du modèle iFHMV à celui des modèles compétiteurs. Pour calculer la prévision de la variable d'intérêt $x_t = r_t^2$ à l'horizon d'un jour pour le modèle iFHMV, nous utilisons l'estimé suivant :

$$\begin{aligned}\hat{x}_{t+1} &= \mathbb{E}(x_{t+1} \mid r_{1:t}, c_{1:\infty}, p_{1:\infty}) \\ &\approx \sum_{m=1}^M \left[\mathbb{E} \left(x_{t+1} \mid \mathcal{S}_{t+1,1:N_+^{(m)}}, c_{1:N_+^{(m)}}, p_{1:N_+^{(m)}} \right) \mathbb{P} \left(\mathcal{S}_{t+1,1:N_+^{(m)}} \mid r_{1:t}, c_{1:N_+^{(m)}}, p_{1:N_+^{(m)}} \right) \right],\end{aligned}$$

où M désigne le nombre d'itérations effectuées, $N_+^{(m)}$ représente le nombre de composantes inféré à l'itération m et $\mathcal{S}_{t+1,1:N_+^{(m)}} = \left(s_{t+1}^{(i)} \right)_{i=1, \dots, N_+^{(m)}}$ représente la $(t+1)$ -ième ligne de la matrice binaire contenant les états des $N_+^{(m)}$ premières chaînes de Markov. Les écarts quadratique et absolu moyens de prévision, ainsi que la quasi-vraisemblance sont respectivement définis par

$$\begin{aligned}\text{RMSFE} &= \sqrt{\frac{1}{T} \sum_{t=0}^{T-1} (\hat{x}_{t+1} - x_{t+1})^2}, \\ \text{MAFE} &= \frac{1}{T} \sum_{t=0}^{T-1} |\hat{x}_{t+1} - x_{t+1}|, \\ \text{QLIK} &= \frac{1}{T} \sum_{t=0}^{T-1} \left[\log(\hat{x}_{t+1}) + \frac{x_{t+1}}{\hat{x}_{t+1}} \right],\end{aligned}$$

où $t = 0$ correspond à la date finale de l'échantillon d'estimation (*in-sample*) et $t = T$ représente la date finale de l'échantillon de prévision (*out-of-sample*). Ces trois mesures de performance liées à la prévision (l'écart quadratique moyen de prévision, l'écart absolu moyen de prévision et la quasi-vraisemblance) sont décroissantes. En d'autres termes, une faible valeur du critère indique une meilleure performance.

Le tableau 4.4 présente les RMSFE, MAFE et QLIK des modèles de référence (iHMM et FHMV) et du modèle iFHMV. Pour chaque série financière et chaque mesure de performance, les valeurs les plus faibles (désignant le meilleur modèle) sont en gras. La signification statistique des différences entre les mesures de performance est évaluée par le test de Diebold-Mariano (Diebold et Mariano, 2002). Pour ce test, nous utilisons le modèle FHMV comme modèle de référence.

D'après ce tableau 4.4, le modèle iFHMV produit toujours un critère plus faible que le modèle iHMM. En le comparant avec le modèle FHMV, dans l'ensemble, le modèle iFHMV produit un critère plus faible (10 cas sur 12). Plus particulièrement, le modèle iFHMV semble produire un critère plus faible que le modèle FHMV pour les données d'indices financiers. Par exemple, les RMSFE, MAFE ET QLIK produits par le modèle iFHMV sont tous plus faibles que ceux produits par le modèle FHMV sur les données de NASDAQ 100 et FTSE 100. De plus, les RMSFE et MAFE produits par le modèle iFHMV sont statistiquement (au seuil de 10%) plus faibles que ceux produits par le modèle FHMV sur les données du NASDAQ 100.

En définitive, contrairement à l'analyse sur les données d'estimation (*in-sample*), le modèle iFHMV produit de meilleures performances par rapport au modèle iHMM sur les données de prévisions (*out-of-sample*).

TABLE 4.4 – Performance de prévision de volatilité

| Modèles | RMSFE | MAFE | QLIK |
|-------------------|-------------------------|--------------|-------------------------|
| NASDAQ 100 | | | |
| FHMV | 2.11 | 1.13 | 0.79 |
| iHMM | 2.29 | 1.29 | 0.80 |
| iFHMV | 2.10* | 1.12* | 0.79 |
| FTSE 100 | | | |
| FHMV | 1.08 | 0.61 | 0.36 |
| iHMM | 1.09 | 0.65 | 0.33 |
| iFHMV | 1.07[†] | 0.60 | 0.32[‡] |
| GES | | | |
| FHMV | 2.96 | 1.80 | 1.48 |
| iHMM | 3.33 | 2.02 | 1.49 |
| iFHMV | 2.96 | 1.85 | 1.46 |
| CS | | | |
| FHMV | 8.95 | 5.00 | 2.39 |
| iHMM | 11.06 | 6.14 | 2.42 |
| iFHMV | 8.78 | 5.10 | 2.38 |

Les symboles *, [†] et [‡] signifient respectivement que le critère est significativement inférieur à celui du modèle de référence (FHMV) aux seuils de 10%, 5% et 1% en utilisant le test de Diebold-Mariano. Les valeurs les plus faibles sont en gras.

4.4.2.3. Nombre de composantes inféré (*out-of-sample*)

Le modèle iFHMV a l’avantage de prendre en compte l’incertitude reliée au nombre de composantes dans le modèle factoriel à chaîne de Markov, aussi bien pour l’estimation que pour la prévision. Nous présentons, à travers la figure 4.6 de cette section, le nombre moyen de composantes inféré pour chaque date de l’échantillon de prévision (*out-of-sample*).

La figure 4.6 montre que la dimension du modèle est relativement stable pour les actions GES et CS et présente des sauts pour les indices NASDAQ 100 et FTSE 100. Pour le NASDAQ 100, la dimension du modèle présente des bonds à différentes dates telles qu’en mars 2018, en juin 2018, en septembre 2018 et en décembre 2018. Le FTSE 100 comporte également des sauts aux dates septembre 2018, janvier 2019 et mars 2019. Remarquons que les sauts apparaissent en général à des moments où les rendements sont plus volatils ou moins volatils. Cela explique que le modèle prend en compte le besoin de composantes supplémentaires permettant de capturer une volatilité supplémentaire.

4.5. Discussion

Les modèles factoriels sont très performants pour modéliser les données financières. Un modèle récent et performant de cette classe de modèles est le modèle FHMV proposé par [Augustyniak et al. \(2019\)](#). Ces modèles factoriels supposent que le processus de variance est lié à une chaîne de Markov de grande dimension, qui est obtenue par le produit d’un nombre fini de chaînes de Markov à deux ou plusieurs états. Dans ces modèles, le nombre de chaînes de Markov sous-jacentes, appelées composantes, est fixé *ex ante* ; ce nombre est

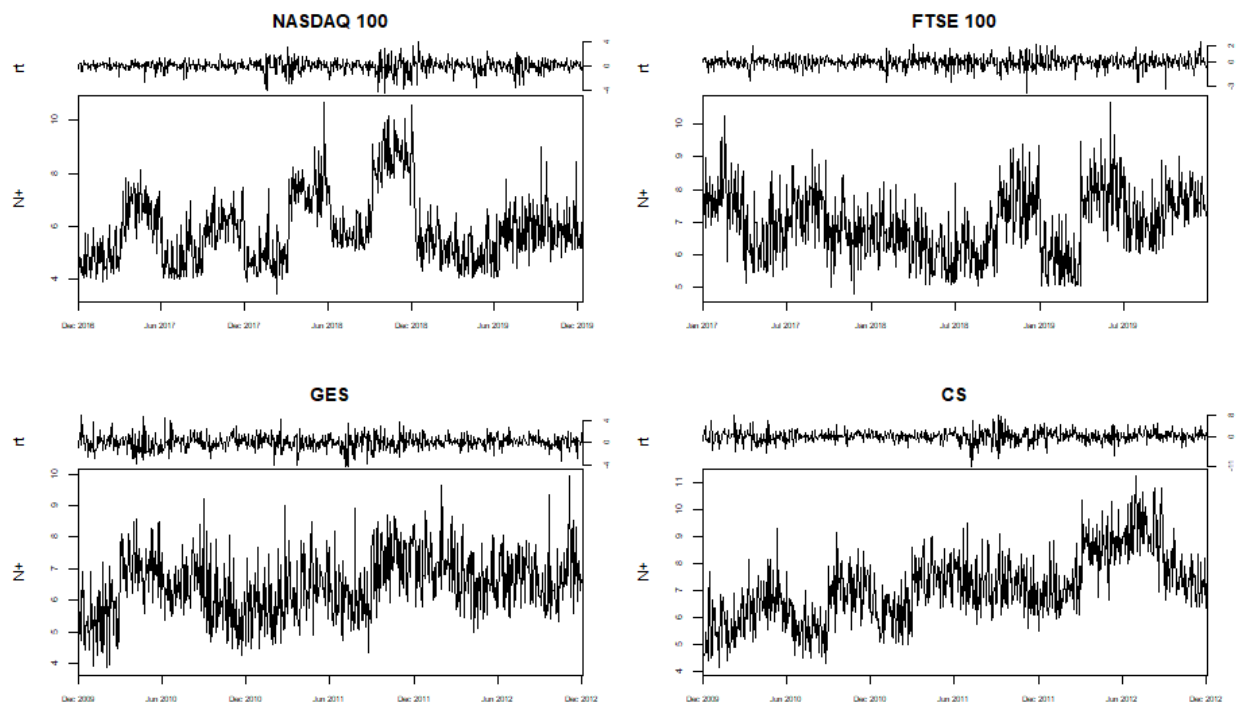


FIGURE 4.6 – Nombre moyen de composantes N_+ inféré pour la prévision sur les 756 jours boursiers. La première ligne présente les graphiques pour les indices financiers NASDAQ 100 et FTSE 100 alors que la deuxième ligne présente les graphiques pour les actions GES et CS. La série de rendements financiers de chaque série est représentée au-dessus du chaque nombre de composantes.

fixé à 10 dans le cas du modèle FHMV. Cependant, des études montrent qu'on pénalise la log-vraisemblance en fixant un nombre de composantes trop faible ou trop élevé. De plus, fixer le nombre de composantes suppose qu'il n'existe aucune incertitude relative à ce dernier. Le présent chapitre étend le modèle FHMV à un cadre où le nombre de composantes est infini, permettant ainsi de tenir compte de l'incertitude liée au nombre de composantes. Nous avons aussi appliqué ce modèle, nommé iFHMV (*infinite Factorial Hidden Volatility*), pour modéliser les rendements de deux indices financiers et de deux actions.

Les résultats empiriques montrent qu'il existe effectivement une incertitude liée au nombre de composantes nécessaires à la modélisation des indices financiers. En effet, l'estimation du modèle iFHMV n'infère pas une valeur particulière pour le nombre de composantes nécessaires à la modélisation de ces données financières. Mieux, le modèle iFHMV prend en compte l'incertitude liée à ce nombre de composantes et se montre plus performant que ses compétiteurs en terme de prévision de la volatilité. Outre la grande flexibilité du modèle iFHMV, une possible explication de sa performance est qu'il permet de capter la tendance des données financières à changer abruptement (les sauts).

Les rendements d'indices financiers présentent, en général, un effet levier. Une extension directe du modèle iFHMV nécessiterait d'intégrer une structure supplémentaire afin de capter cet effet levier. Cette structure permettrait ainsi de faire de meilleures prévisions de rendements financiers. De plus, intégrer les

variances réalisées dans la modélisation permettrait d'améliorer les performances d'estimation et de prévision du modèle iFHMV.

Conclusion générale

Cette thèse traite de la modélisation des données financières par les modèles à chaîne de Markov cachée de haute dimension. Premièrement, elle se base sur une approche fréquentiste et propose un modèle factoriel à chaîne de Markov cachée, le modèle MDSV (*Multifractal Discrete Stochastic Volatility*). Deuxièmement, elle propose un modèle bayésien non-paramétrique à partir de la classe des modèles factoriels à chaîne de Markov cachée, le modèle iFHMV (*infinite Factorial Hidden Markov Volatility*).

Le premier chapitre de cette thèse construit la littérature de modélisation des données financières tout en présentant les processus de données financières utilisées (rendements financiers et variances réalisées) et les modèles financiers populaires (modèles GARCH, modèles à volatilité stochastique, modèles MEM, modèles HAR). Ce chapitre introduit aussi le cadre de modélisation des modèles à chaîne de Markov cachée tout en présentant leur procédure d'estimation et de prévision. Quelques modèles à chaîne de Markov cachée de la littérature (MSM de [Calvet et Fisher, 2004](#), CDRS de [Fleming et Kirby, 2013](#), FHMV de [Augustyniak et al., 2019](#) et DSARV de [Cordis et Kirby, 2014](#)) sont également présentés au chapitre 1.

Dans le chapitre 2, nous avons défini formellement le cadre de modélisation factorielle et indiqué que les modèles MSM, CDRS, FHMV et DSARV adhèrent à ce cadre. Il s'agit de modèles financiers dont le processus de volatilité est obtenu à partir du produit de N chaînes de Markov appelées composantes. Nous avons aussi étudié les propriétés théoriques associées à ce cadre et prouvé que les modèles adhérant à ce cadre partagent des similitudes avec les modèles à volatilité stochastique. Nous introduisons également le processus MDSV qui généralise les modèles factoriels en autorisant les composantes à avoir plus de 2 états. La paramétrisation particulière du processus MDSV lui accorde de la parcimonie (seulement 5 paramètres) et suffisamment de flexibilité. L'analyse des propriétés théoriques du processus MDSV indique que la fonction d'autocorrélation théorique dépend du nombre N de composantes et du nombre K d'états de chaque composante. Cela permet au processus MDSV de refléter différentes allures des fonctions d'autocorrélation et en particulier celles qui décroissent lentement. L'étude de la distribution du temps de séjour des états du processus MDSV révèle que le temps passé par le processus MDSV dans certains de ses états n'est pas distribué selon une loi géométrique. Cette propriété permet de créer un effet semi-markovien qui produit une décroissance plus lente de la fonction d'autocorrélation.

Le processus MDSV peut être utilisé pour modéliser, séparément ou conjointement, des données de rendements financiers et de variances réalisées. Le chapitre 2 propose ainsi un cadre pour relier les processus

de rendements financiers et/ou de variances réalisées au processus MDSV et déduit des propriétés théoriques (distribution conditionnelle, moments, fonction d'autocorrélation, etc.) concernant ces processus. Le chapitre 2 développe également les procédures d'estimation et de prévision correspondant à ce cadre, puis propose un moyen de tenir compte de l'effet levier dans le modèle. Bien que ce moyen de tenir compte de l'effet levier soit élégant, il ne permet pas de calculer les moments du modèle et donc de faire des prévisions d'horizon $h > 1$ analytiquement. Un champ de recherche futur, que j'explore actuellement, est de proposer un processus d'effet levier avec lequel on peut calculer les moments.

Dans le chapitre 3, nous effectuons une analyse empirique du modèle MDSV introduit au chapitre 2. Les données de rendements financiers et de variances réalisées d'indices financiers sont utilisées à cet effet. Ce chapitre expose ainsi les performances d'ajustement, de prévisions de volatilité et de gestion du risque de différentes spécifications du modèle MDSV tout en les comparant à des modèles de référence. Les spécifications univariées (modélisant uniquement les rendements financiers ou les variances réalisées) du modèle MDSV produisent des performances d'ajustement et de prévision similaires à celles des modèles de référence. La spécification jointe, quant à elle, domine significativement les modèles de référence, aussi bien en termes d'ajustement que de prévision de volatilité et de log-vraisemblance prédictive, comme le montre l'analyse sur les données de 31 indices financiers. À travers le modèle MDSV, nous avons étudié empiriquement l'avantage comparatif d'un modèle joint par rapport à un modèle univariée et avons conclu que la forme jointe est plus performante que la forme univariée.

Empiriquement, nous avons constaté que la performance d'ajustement du modèle MDSV augmente avec le nombre de composantes jusqu'à une certaine valeur, puis se met à diminuer légèrement. L'estimation du modèle MDSV sur les données simulées, au chapitre 2, révèle également qu'un trop faible nombre ou un nombre trop élevé de composantes réduit les performances d'estimation du modèle. De plus, en fixant, un nombre de composantes, on suppose qu'il n'existe aucune incertitude liée à ce nombre. Le chapitre 4 propose un modèle, le modèle iFHMV, autorisant les données à déterminer le nombre adéquat de composantes nécessaires à leur modélisation. Le modèle iFHMV est un modèle bayésien non-paramétrique se basant sur le processus IBP (*Indian Buffet Process*). Nous développons un algorithme d'inférence pour le modèle iFHMV et l'appliquons pour l'estimation et la prévision de volatilité d'indices financiers et d'actions. Les résultats d'estimation montrent qu'il existe effectivement une incertitude liée au nombre de composantes nécessaires à la modélisation des données financières. En terme de prévision, une comparaison du modèle iFHMV avec des modèles de référence montre que le modèle iFHMV présente les meilleures performances. L'analyse du nombre de composantes inféré montre aussi que le modèle iFHMV prend en compte le besoin de composantes supplémentaires. Bien que le modèle iFHMV présente une bonne performance de prévision, il n'intègre pas l'effet levier. De plus, le modèle iFHMV est appliqué uniquement aux données de rendements financiers. Des extensions possibles de ce modèle pourraient prendre en compte l'effet levier afin de proposer une approche de modélisation jointe des données de rendements financiers et de variances réalisées.

Cette thèse offre un cadre formel de modélisation des données financières par l'entremise des modèles factoriels à chaîne de Markov cachée. Cette contribution ouvre la voie à plusieurs analyses sur les données financières. Premièrement, le cadre de modélisation proposé dans cette thèse permet d'étendre l'analyse des données financières à une dimension multivariée. Jusqu'à présent, mes recherches ont porté sur l'étude d'un seul indice financier. Cependant, il existe une certaine dépendance entre l'évolution des rendements de la plupart des indices financiers. Les modèles factoriels à chaîne de Markov cachée offrent une possibilité de modéliser simultanément plusieurs indices financiers en les reliant par un processus latent commun. Ce dernier permettra ainsi de détecter et étudier les similarités existant entre différentes données financières. Deuxièmement, les bons résultats en termes d'estimation et de prévision obtenus avec les modèles factoriels suggèrent des applications pour la tarification d'option et la couverture de risque. Par exemple, puisque les modèles factoriels à chaîne de Markov cachée captent mieux la haute persistance et la haute variabilité dans la volatilité des indices financiers, ils devraient permettre de tarifier plus précisément des options avec une longue maturité.

Bibliographie

- Sassan ALIZADEH, Michael W BRANDT et Francis X DIEBOLD : Range-based estimation of stochastic volatility models. *The Journal of Finance*, 57(3):1047–1091, 2002.
- Torben G ANDERSEN : Stochastic autoregressive volatility : a framework for volatility modeling. *Mathematical finance*, 4:75–102, 1994.
- Torben G. ANDERSEN et Tim BOLLERSLEV : Heterogeneous information arrivals and return volatility dynamics : Uncovering the long-run in high frequency returns. *The Journal of Finance*, 52(3):975–1005, 1997. ISSN 1540-6261. URL <http://dx.doi.org/10.1111/j.1540-6261.1997.tb02722.x>.
- Torben G ANDERSEN, Tim BOLLERSLEV et Francis X DIEBOLD : Roughing it up : Including jump components in the measurement, modeling, and forecasting of return volatility. *The review of economics and statistics*, 89(4):701–720, 2007.
- Torben G ANDERSEN, Tim BOLLERSLEV, Francis X DIEBOLD et Heiko EBENS : The distribution of realized stock return volatility. *Journal of financial economics*, 61(1):43–76, 2001.
- Andrew ANG et Geert BEKAERT : Regime switches in interest rates. *Journal of Business & Economic Statistics*, 20(2):163–182, 2002.
- Maciej AUGUSTYNIAK, Luc BAUWENS et Arnaud DUFAYS : A new approach to volatility modeling : the factorial hidden markov volatility model. *Journal of Business & Economic Statistics*, 37(4):696–709, 2019.
- Richard T BAILLIE : Long memory processes and fractional integration in econometrics. *Journal of econometrics*, 73(1):5–59, 1996.
- Richard T BAILLIE, Tim BOLLERSLEV et Hans Ole MIKKELSEN : Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 74(1):3–30, 1996.
- David S BATES : Jumps and stochastic volatility : Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107, 1996.
- Leonard E. BAUM, Ted PETRIE, George SOULES et Norman WEISS : A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, 41:164–171, 1970. ISSN 0003-4851.
- Luc BAUWENS, Arnaud DUFAYS et Jeroen VK ROMBOUTS : Marginal likelihood for Markov-switching and change-point GARCH models. *Journal of Econometrics*, 178:508–522, 2014.

- Luc BAUWENS, Christian M HAFNER et Sébastien LAURENT : *Handbook of volatility models and their applications*, volume 3. John Wiley & Sons, 2012.
- Tim BOLLERSLEV : Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- Christian T BROWNLEES, Fabrizio CIPOLLINI et Giampiero M GALLO : Multiplicative error models. *Available at SSRN 1852285*, 2011.
- Jan BULLA et Ingo BULLA : Stylized facts of financial time series and hidden semi-Markov models. *Computational Statistics & Data Analysis*, 51(4):2192–2209, 2006.
- Laurent CALVET et Adlai FISHER : Forecasting multifractal volatility. *Journal of econometrics*, 105(1):27–58, 2001.
- Laurent E CALVET et Adlai J FISHER : How to forecast long-run volatility : regime switching and the estimation of multifractal processes. *Journal of Financial Econometrics*, 2(1):49–83, 2004.
- Sean D CAMPBELL : A review of backtesting and backtesting procedures. 2005.
- Leopoldo CATANIA : A stochastic volatility model with a general leverage specification. *Journal of Business & Economic Statistics*, pages 1–23, 2020.
- Liyuan CHEN, Paola ZERILLI et Christopher F BAUM : Leverage effects and stochastic volatility in spot oil returns : A bayesian approach with var and cvar applications. *Energy Economics*, 79:111–129, 2019.
- Siddhartha CHIB : Calculating posterior distributions and modal estimates in markov mixture models. *Journal of Econometrics*, 75(1):79–97, 1996.
- Peter CHRISTOFFERSEN et Denis PELLETIER : Backtesting value-at-risk : A duration-based approach. *Journal of Financial Econometrics*, 2(1):84–108, 2004.
- Peter F CHRISTOFFERSEN : Evaluating interval forecasts. *International economic review*, pages 841–862, 1998.
- Fabrizio CIPOLLINI, Robert F ENGLE et Giampiero M GALLO : Semiparametric vector mem. *Journal of Applied Econometrics*, 28(7):1067–1086, 2013.
- Fabrizio CIPOLLINI, Giampiero M GALLO, Edoardo OTRANTO *et al.* : Realized volatility forecasting : Robustness to measurement errors. Rapport technique, Università degli Studi di Firenze, Dipartimento di Statistica, Informatica . . . , 2019a.
- Fabrizio CIPOLLINI, Giampiero M GALLO, Alessandro PALANDRI *et al.* : Realized variance modeling : decoupling forecasting from estimation. Rapport technique, Università degli Studi di Firenze, Dipartimento di Statistica, Informatica . . . , 2019b.
- Jerome B COHEN, Fischer BLACK et Myron SCHOLES : The valuation of option contracts and a test of market efficiency. *The Journal of Finance*, 27(2):399–417, 1972.
- Rama CONT : Empirical properties of asset returns : stylized facts and statistical issues. *Quantitative Finance*, 1:223 – 236, 2001.

- Adriana S CORDIS et Chris KIRBY : Discrete stochastic autoregressive volatility. *Journal of Banking & Finance*, 43:160–178, 2014.
- Fulvio CORSI : A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196, 2009.
- Fulvio CORSI, Davide PIRINO et Roberto RENO : Volatility forecasting : The jumps do matter. 2009.
- Fulvio CORSI et Roberto RENÒ : Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business & Economic Statistics*, 30(3):368–380, 2012.
- Giovanni DE LUCA, Giampiero M GALLO *et al.* : A time-varying mixing multiplicative error model for realized volatility. *Econometrics Working Papers Archive wp2010–03*, Università degli Studi di Firenze, Dipartimento di Statistica “G. Parenti”, 2010.
- Francis X DIEBOLD et Robert S MARIANO : Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144, 2002.
- Zhuanxin DING, Clive WJ GRANGER et Robert F ENGLE : A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106, 1993.
- Arnaud DUFAYS : Infinite-state markov-switching for dynamic volatility. *Jnl of Financial Econometrics*, 14(2):418–460, 2015.
- Darrell DUFFIE et Kenneth J SINGLETON : Simulated moments estimation of markov models of asset prices. Rapport technique, National Bureau of Economic Research, 1990.
- Heiko EBENS *et al.* : Realized stock volatility. *Department of Economics, Johns Hopkins University*, 1999.
- Robert ENGLE : New frontiers for ARCH models. *Journal of Applied Econometrics*, 17(5):425–446, 2002.
- Robert F ENGLE : Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica : Journal of the Econometric Society*, pages 987–1007, 1982.
- Robert F ENGLE et Giampiero M GALLO : A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131(1-2):3–27, 2006.
- Robert F ENGLE, Giampiero M GALLO et Margherita VELUCCHI : Volatility spillovers in east asian financial markets : a mem-based approach. *Review of Economics and Statistics*, 94(1):222–223, 2012.
- Robert F ENGLE et Victor K NG : Measuring and testing the impact of news on volatility. *The journal of finance*, 48(5):1749–1778, 1993.
- Jeff FLEMING et Chris KIRBY : Component-driven regime-switching volatility. *Journal of Financial Econometrics*, 11(2):263–301, 2013.
- Moshe FRIDMAN et Lawrence HARRIS : A maximum likelihood approach for non-gaussian stochastic volatility models. *Journal of Business & Economic Statistics*, 16(3):284–291, 1998.
- Jurgen V GAEL, Yee W TEH et Zoubin GHARAMANI : The infinite factorial hidden markov model. *In Advances in Neural Information Processing Systems*, pages 1697–1704, 2009.

- Giampiero M. GALLO et Edoardo OTRANTO : Forecasting realized volatility with changing average levels. *International Journal of Forecasting*, 31(3):620–634, 2015. ISSN 0169-2070. URL <http://www.sciencedirect.com/science/article/pii/S0169207014001812>.
- Zoubin GHAHRAMANI et Thomas L GRIFFITHS : Infinite latent feature models and the indian buffet process. *In Advances in neural information processing systems*, pages 475–482, 2006.
- Zoubin GHAHRAMANI et Michael I. JORDAN : Factorial hidden Markov models. *Machine Learning*, 29(2):245–273, 1997. ISSN 1573-0565. URL <http://dx.doi.org/10.1023/A:1007425814087>.
- Lawrence R GLOSTEN, Ravi JAGANNATHAN et David E RUNKLE : On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, 48(5):1779–1801, 1993.
- Clive WJ GRANGER et Roselyne JOYEUX : An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis*, 1(1):15–29, 1980.
- James HAMILTON : D.(1994), time series analysis, 1994.
- James D HAMILTON : A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica : Journal of the Econometric Society*, pages 357–384, 1989.
- James D HAMILTON : Analysis of time series subject to changes in regime. *Journal of econometrics*, 45(1-2):39–70, 1990.
- Hyojin HAN, Stanislav KHRAPOV et Eric RENAULT : The leverage effect puzzle revisited : Identification in discrete time. *Journal of Econometrics*, 217(2):230 – 258, 2020. ISSN 0304-4076. URL <http://www.sciencedirect.com/science/article/pii/S0304407619302490>. Nonlinear Financial Econometrics.
- Peter Reinhard HANSEN et Zhuo HUANG : Exponential garch modeling with realized measures of volatility. *Journal of Business & Economic Statistics*, 34(2):269–287, 2016.
- Peter Reinhard HANSEN, Zhuo HUANG et Howard Howan SHEK : Realized garch : a joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906, 2012.
- Mary HARDY : *Investment guarantees : modeling and risk management for equity-linked life insurance*, volume 215. John Wiley & Sons, 2003.
- Brian M HARTMAN et Matthew J HEATON : Accounting for regime and parameter uncertainty in regime-switching models. *Insurance : Mathematics and Economics*, 49(3):429–437, 2011.
- Steven L HESTON : A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343, 1993.
- J. R. M. HOSKING : Fractional differencing. *Biometrika*, 68(1):165–176, 1981. ISSN 00063444. URL <http://www.jstor.org/stable/2335817>.
- Shirley HUANG, Qianqiu LIU et YU JUN : Realized daily variance of s&p 500 cash index : a revaluation of stylized facts. *Annals of Economics and Finance*, 8:33, 2007.
- John HULL et Alan WHITE : The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300, 1987.

- Philippe JORION : Value at risk. 2000.
- Tomáš KONDERLA, Václav KLEPÁČ *et al.* : Using hmm approach for assessing quality of value at risk estimation : Evidence from pse listed company. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 65(5):1687–1694, 2017.
- Keith KUESTER, Stefan MITTNIK et Marc S PAOLELLA : Value-at-risk prediction : A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89, 2006.
- Paul KUPIEC : Techniques for verifying the accuracy of risk measurement models. *The J. of Derivatives*, 3(2), 1995.
- Markku LANGE : A mixture multiplicative error model for realized volatility. *Journal of Financial Econometrics*, 4(4):594–616, 2006. URL <http://jfec.oxfordjournals.org/content/4/4/594.abstract>.
- Jia LIU et John M MAHEU : Improving markov switching models using realized variance. *Journal of Applied Econometrics*, 33(3):297–318, 2018.
- Leandro MACIEL : Cryptocurrencies value-at-risk and expected shortfall : Do regime-switching volatility models improve forecasting? *International Journal of Finance & Economics*, 2020.
- John M MAHEU et Thomas H MCCURDY : Do high-frequency measures of volatility improve forecasts of return distributions? *Journal of Econometrics*, 160(1):69–76, 2011.
- John M MAHEU et Qiao YANG : An infinite hidden markov model for short-term interest rates. *Journal of Empirical Finance*, 38:202–220, 2016.
- Antonello MARUOTTI, Antonio PUNZO et Luca BAGNATO : Hidden markov and semi-markov models with multivariate leptokurtic-normal components for robust modeling of daily returns series. *Journal of Financial Econometrics*, 17(1):91–117, 2019.
- Thomas H MCCURDY et Ieuan G MORGAN : Tests of the martingale hypothesis for foreign currency futures with time-varying volatility. *International Journal of Forecasting*, 3(1):131–148, 1987.
- Daniel B NELSON : Conditional heteroskedasticity in asset returns : A new approach. *Econometrica : Journal of the Econometric Society*, pages 347–370, 1991.
- Andrew J PATTON : Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256, 2011.
- Christian P ROBERT, Tobias RYDÉN et David M TITTERINGTON : Convergence controls for mcmc algorithms, with applications to hidden markov chains. *Journal of Statistical Computation and Simulation*, 64(4):327–355, 1999.
- Gerardo RUBINO et Bruno SERICOLA : Sojourn times in finite markov processes. *Journal of Applied Probability*, 26(4):744–756, 1989.
- Tobias RYDÉN, Timo TERÄSVIRTA et Stefan ÅSBRINK : Stylized facts of daily return series and the hidden markov model. *Journal of applied econometrics*, 13(3):217–244, 1998.
- Steven L SCOTT : Bayesian methods for hidden markov models : Recursive computing in the 21st century. *Journal of the American Statistical Association*, 97(457):337–351, 2002.

- Eugene SENETA : *Non-negative matrices and Markov chains*. Springer Science & Business Media, 2006.
- Makoto TAKAHASHI, Yasuhiro OMORI et Toshiaki WATANABE : Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics & Data Analysis*, 53(6):2404–2426, 2009.
- Makoto TAKAHASHI, Yasuhiro OMORI et Toshiaki WATANABE : News impact curve for stochastic volatility models. *Economics Letters*, 120(1):130–134, 2013.
- Makoto TAKAHASHI, Toshiaki WATANABE, Yasuhiro OMORI *et al.* : Forecasting daily volatility of stock price index using daily returns and realized volatility. Rapport technique, 2021.
- Yee Whye TEH, Dilan GRÜR et Zoubin GHAHRAMANI : Stick-breaking construction for the indian buffet process. *In Artificial Intelligence and Statistics*, pages 556–563, 2007.
- Isabel VALERA, Francisco JR RUIZ et Fernando PEREZ-CRUZ : Infinite factorial unbounded-state hidden markov model. *IEEE transactions on pattern analysis and machine intelligence*, 38(9):1816–1828, 2015.
- Qi WANG et Zerong WANG : Vix futures and its closed-form pricing through an affine garch model with realized variance. *Journal of Futures Markets*, 41(1):135–156, 2021.
- Walter ZUCCHINI et Iain L. MACDONALD : *Hidden Markov models for time series*, volume 110 de *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2009. ISBN 978-1-58488-573-3. URL <http://dx.doi.org/10.1201/9781420010893>. An introduction using R.

Annexe A

Annexe relative au chapitre 2

A.1. Preuve du théorème 2.1.1

Théorème. Soit $N \geq 2$ et $\{C_t^{(1)}\}, \{C_t^{(2)}\}, \dots, \{C_t^{(N)}\}$ des chaînes de Markov discrètes, homogènes et indépendantes à espaces d'états finis respectifs $E^{(1)}, E^{(2)}, \dots, E^{(N)}$ et de matrices de transition respectives $\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_N$.

Le processus $\{M_t\}$, défini pour chaque t par $M_t = (C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)})$, est une chaîne de Markov discrète et homogène à espace d'états fini $E_M = E^{(1)} \times E^{(2)} \times \dots \times E^{(N)}$ de matrice de transition $\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \dots \otimes \mathbf{P}_N$, où le symbole \otimes représente le produit de Kronecker et \times représente le produit des ensembles.

DÉMONSTRATION. Montrons d'abord que $\{M_t\}$ est une chaîne de Markov. Cela revient à démontrer que pour tout t , on a l'égalité

$$\mathbb{P}[M_t = m_t \mid M_{t-1} = m_{t-1}, \dots, M_0 = m_0] = \mathbb{P}[M_t = m_t \mid M_{t-1} = m_{t-1}].$$

Pour tout t , on a $M_t = (C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)})$ et $m_t = (c_t^{(1)}, c_t^{(2)}, \dots, c_t^{(N)})$ où $c_t^{(i)} \in E^{(i)}$ pour tout $i = 1, 2, \dots, N$. Puisque les chaînes de Markov $\{C_t^{(1)}\}, \{C_t^{(2)}\}, \dots, \{C_t^{(N)}\}$ sont indépendantes, on a successivement :

$$\begin{aligned} \mathbb{P}[M_t = m_t \mid M_{t-1} = m_{t-1}, \dots, M_0 = m_0] &= \mathbb{P}\left[C_t^{(1)} = c_t^{(1)}, \dots, C_t^{(N)} = c_t^{(N)} \mid M_{t-1} = m_{t-1}, \dots, M_0 = m_0\right], \\ &= \prod_{i=1}^N \mathbb{P}\left[C_t^{(i)} = c_t^{(i)} \mid M_{t-1} = m_{t-1}, \dots, M_0 = m_0\right], \\ &= \prod_{i=1}^N \mathbb{P}\left[C_t^{(i)} = c_t^{(i)} \mid C_{t-1}^{(i)} = c_{t-1}^{(i)}, \dots, C_0^{(i)} = c_0^{(i)}\right], \\ &= \prod_{i=1}^N \mathbb{P}\left[C_t^{(i)} = c_t^{(i)} \mid C_{t-1}^{(i)} = c_{t-1}^{(i)}\right]. \end{aligned}$$

La dernière égalité est obtenue puisque, par hypothèse, les processus $\{C_t^{(i)}\}$ pour $i = 1, 2, \dots, N$ sont des chaînes de Markov. Puisque ces chaînes sont indépendantes, on a :

$$\begin{aligned} \mathbb{P}[M_t = m_t \mid M_{t-1} = m_{t-1}, \dots, M_0 = m_0] &= \prod_{i=1}^N \mathbb{P}\left[C_t^{(i)} = c_t^{(i)} \mid C_{t-1}^{(1)} = c_{t-1}^{(1)}, C_{t-1}^{(2)} = c_{t-1}^{(2)}, \dots, C_{t-1}^{(N)} = c_{t-1}^{(N)}\right] . \\ &= \prod_{i=1}^N \mathbb{P}\left[C_t^{(i)} = c_t^{(i)} \mid M_{t-1} = m_{t-1}\right] . \\ &= \mathbb{P}\left[C_t^{(1)} = c_t^{(1)}, \dots, C_t^{(N)} = c_t^{(N)} \mid M_{t-1} = m_{t-1}\right] , \\ &= \mathbb{P}[M_t = m_t \mid M_{t-1} = m_{t-1}] . \end{aligned}$$

Ce qui prouve que le processus $\{M_t\}$ est une chaîne de Markov. La preuve de l'homogénéité de la chaîne $\{M_t\}$ sera faite par induction.

Début ($N = 2$) : Posons $E^{(1)} = \{u_1, u_2, \dots, u_r\}$ et $E^{(2)} = \{v_1, v_2, \dots, v_s\}$. Par définition, l'espace d'états du processus $\{M_t\}$ défini pour tout t par $M_t = (C_t^{(1)}, C_t^{(2)})$ est $E_M = E^{(1)} \times E^{(2)}$. Soit $m_{ij} = (u_i, v_j)$ et $m_{kl} = (u_k, v_l)$, où $i, k \in \{1, \dots, r\}$ et $j, l \in \{1, \dots, s\}$. Notre objectif est de montrer que la chaîne $\{M_t\}$ est homogène c'est-à-dire

$$\mathbb{P}[M_{t+1} = m_{ij} \mid M_t = m_{kl}] = \mathbb{P}[M_t = m_{ij} \mid M_{t-1} = m_{kl}] .$$

On a :

$$\begin{aligned} \mathbb{P}[M_{t+1} = m_{ij} \mid M_t = m_{kl}] &= \mathbb{P}\left[\left(C_{t+1}^{(1)}, C_{t+1}^{(2)}\right) = (u_i, v_j) \mid \left(C_t^{(1)}, C_t^{(2)}\right) = (u_k, v_l)\right] , \\ &= \mathbb{P}\left[C_{t+1}^{(1)} = u_i \mid \left(C_t^{(1)}, C_t^{(2)}\right) = (u_k, v_l)\right] \mathbb{P}\left[C_{t+1}^{(2)} = v_j \mid \left(C_t^{(1)}, C_t^{(2)}\right) = (u_k, v_l)\right] , \\ &= \mathbb{P}\left[C_{t+1}^{(1)} = u_i \mid C_t^{(1)} = u_k\right] \mathbb{P}\left[C_{t+1}^{(2)} = v_j \mid C_t^{(2)} = v_l\right] \end{aligned} \quad (\text{A.1.1})$$

par indépendance de $\{C_t^{(1)}\}$ et $\{C_t^{(2)}\}$. En utilisant l'homogénéité des chaînes, nous pouvons alors écrire

$$\begin{aligned} \mathbb{P}[M_{t+1} = m_{ij} \mid M_t = m_{kl}] &= \mathbb{P}\left[C_t^{(1)} = u_i \mid C_{t-1}^{(1)} = u_k\right] \mathbb{P}\left[C_t^{(2)} = v_j \mid C_{t-1}^{(2)} = v_l\right] , \\ &= \mathbb{P}\left[C_t^{(1)} = u_i \mid \left(C_{t-1}^{(1)}, C_{t-1}^{(2)}\right) = (u_k, v_l)\right] \mathbb{P}\left[C_t^{(2)} = v_j \mid \left(C_{t-1}^{(1)}, C_{t-1}^{(2)}\right) = (u_k, v_l)\right] , \\ &= \mathbb{P}\left[\left(C_t^{(1)}, C_t^{(2)}\right) = (u_i, v_j) \mid \left(C_{t-1}^{(1)}, C_{t-1}^{(2)}\right) = (u_k, v_l)\right] , \\ &= \mathbb{P}[M_t = m_{ij} \mid M_{t-1} = m_{kl}] . \end{aligned}$$

Le processus $\{M_t\}$ est donc une chaîne de Markov discrète homogène à espace d'états fini $E_M = E^{(1)} \times E^{(2)}$ et à matrice de transition $\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2$ (équation (A.1.1)).

Induction : Supposons pour N fixé que $M_t = (C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)})$ soit une chaîne de Markov et montrons que pour une nouvelle chaîne de Markov discrète, homogène $\{C_t^{(N+1)}\}$ à espace d'états fini $E^{(N+1)}$, de matrice de transition \mathbf{P}_{N+1} et indépendante des processus $\{C_t^{(1)}\}, \{C_t^{(2)}\}, \dots, \{C_t^{(N)}\}$, le processus $\{N_t\}$ défini pour tout t par $N_t = (C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)}, C_t^{(N+1)})$ est une chaîne de Markov discrète

et homogène à espace d'états fini $E_N = E^{(1)} \times E^{(2)} \times \dots \times E^{(N)} \times E^{(N+1)}$ de matrice de transition $\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \dots \otimes \mathbf{P}_N \otimes \mathbf{P}_{N+1}$.

Notons que $N_t = \left(C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)}, C_t^{(N+1)} \right) = \left(M_t, C_t^{(N+1)} \right)$. Puisque $\{M_t\}$ et $\{C_t^{(N+1)}\}$ sont des chaînes de Markov discrètes, homogènes et à espace d'états de cardinalité finie, alors, par hypothèse d'induction, le processus $\{N_t\}$ est aussi une chaîne de Markov discrète, homogène à espace d'états fini $E_M \times E^{(N+1)} = E^{(1)} \times E^{(2)} \times \dots \times E^{(N)} \times E^{(N+1)}$ et de matrice de transition $\mathbf{P} \otimes \mathbf{P}_{N+1} = \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \dots \otimes \mathbf{P}_N \otimes \mathbf{P}_{N+1}$.

Conclusion : Pour tout $N \geq 2$, le processus $\{M_t\}$, défini pour chaque t par $M_t = \left(C_t^{(1)}, C_t^{(2)}, \dots, C_t^{(N)} \right)$, est une chaîne de Markov discrète et homogène à espace d'états fini $E_M = E^{(1)} \times E^{(2)} \times \dots \times E^{(N)}$ de matrice de transition $\mathbf{P} = \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \dots \otimes \mathbf{P}_N$.

□

A.2. Preuves de la proposition 2.1.1

Proposition. Soit V , la chaîne de Markov définie par l'équation (2.1.1). En posant $h_t^{(i)} = \log \left(C_t^{(i)} \right)$, on obtient que :

(i) Le processus $\{h_t^{(i)}\}$ a une représentation autorégressive d'ordre 1 :

$$h_{t+1}^{(i)} = c_i + \phi^{(i)} h_t^{(i)} + u_{t+1}^{(i)},$$

où $c_i = (1 - \phi^{(i)}) \boldsymbol{\pi}^{(i)'} \log \boldsymbol{\nu}^{(i)}$ et $u_t^{(i)}$ est un processus vérifiant $\mathbb{E} \left[u_t^{(i)} \mid h_{t-1}^{(i)}, \dots, h_1^{(i)} \right] = 0$.

(ii) Le processus V vérifie la relation

$$V_t = \frac{\sigma^2}{\prod_{i=1}^N [\boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)}]} \exp \left(\sum_{i=1}^N h_t^{(i)} \right).$$

DÉMONSTRATION.

(i) Soit $\mathbf{x}_t^{(i)}$, le vecteur $K \times 1$ indiquant l'état de la chaîne (lorsque la chaîne de Markov $C^{(i)}$ est à l'état j au temps t , la j -ième composante de $\mathbf{x}_t^{(i)}$ est 1 et toutes ses autres composantes sont nulles). Posons

$$\boldsymbol{\nu}^{(i)} = \left(\nu_1^{(i)} \quad \nu_2^{(i)} \quad \dots \quad \nu_K^{(i)} \right)'$$

et

$$\iota^{(i)} = \log \left(\boldsymbol{\nu}^{(i)} \right) = \left(\log \left(\nu_1^{(i)} \right) \quad \log \left(\nu_2^{(i)} \right) \quad \dots \quad \log \left(\nu_K^{(i)} \right) \right)'$$

La représentation autorégressive de la chaîne de Markov $C^{(i)}$ est donnée par

$$\begin{aligned} C_t^{(i)} &= \boldsymbol{\nu}^{(i)'} \mathbf{x}_t^{(i)}, \\ \mathbf{x}_{t+1}^{(i)} &= \mathbf{P}^{(i)'} \mathbf{x}_t^{(i)} + \mathbf{e}_{t+1}^{(i)}, \end{aligned}$$

où $\mathbf{e}_t^{(i)}$ est un processus tel que $\mathbb{E} \left[\mathbf{e}_t^{(i)} \mid \mathbf{x}_{t-1}^{(i)}, \dots, \mathbf{x}_1^{(i)} \right] = \mathbf{0}$. De plus,

$$h_t^{(i)} = \log \left(C_t^{(i)} \right) = \iota^{(i)'} \mathbf{x}_t^{(i)}.$$

On obtient successivement

$$\begin{aligned}
h_{t+1}^{(i)} &= \iota^{(i)'} \mathbf{x}_{t+1}^{(i)}, \\
&= \iota^{(i)'} \left[\mathbf{P}^{(i)'} \mathbf{x}_t^{(i)} + \mathbf{e}_{t+1}^{(i)} \right], \\
&= \iota^{(i)'} \left\{ \left[\phi^{(i)} \mathbf{I}_K + (1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'} \right]' \mathbf{x}_t^{(i)} + \mathbf{e}_{t+1}^{(i)} \right\}, \\
&= \iota^{(i)'} \left\{ \phi^{(i)} \mathbf{x}_t^{(i)} + (1 - \phi^{(i)}) \boldsymbol{\pi}^{(i)} \mathbf{1}'_K \mathbf{x}_t^{(i)} + \mathbf{e}_{t+1}^{(i)} \right\}.
\end{aligned}$$

Puisqu'une seule composante de $\mathbf{x}_t^{(i)}$ est égale à 1 et que les autres sont nulles alors $\mathbf{1}'_K \mathbf{x}_t^{(i)} = 1$.

Ainsi,

$$\begin{aligned}
h_{t+1}^{(i)} &= \phi^{(i)} \underbrace{\iota^{(i)'} \mathbf{x}_t^{(i)}}_{h_t^{(i)}} + \underbrace{(1 - \phi^{(i)}) \iota^{(i)'} \boldsymbol{\pi}^{(i)}}_{c_i} + \underbrace{\iota^{(i)'} \mathbf{e}_{t+1}^{(i)}}_{u_{t+1}^{(i)}}, \\
&= c_i + \phi^{(i)} h_t^{(i)} + u_{t+1}^{(i)},
\end{aligned}$$

où $c_i = (1 - \phi^{(i)}) \iota^{(i)'} \boldsymbol{\pi}^{(i)} = (1 - \phi^{(i)}) \boldsymbol{\pi}^{(i)'} \log \boldsymbol{\nu}^{(i)}$ et $u_{t+1}^{(i)} = \iota^{(i)'} \mathbf{e}_{t+1}^{(i)}$. On a aussi $\mathbb{E} \left[u_t^{(i)} \mid h_{t-1}^{(i)}, \dots, h_1^{(i)} \right] = 0$ puisque

$$\begin{aligned}
\mathbb{E} \left[u_t^{(i)} \mid h_{t-1}^{(i)}, \dots, h_1^{(i)} \right] &= \mathbb{E} \left[\iota^{(i)'} \mathbf{e}_t^{(i)} \mid h_{t-1}^{(i)}, \dots, h_1^{(i)} \right], \\
&= \iota^{(i)'} \mathbb{E} \left[\mathbf{e}_t^{(i)} \mid h_{t-1}^{(i)}, \dots, h_1^{(i)} \right], \\
&= \iota^{(i)'} \mathbb{E} \left[\mathbf{e}_t^{(i)} \mid \mathbf{x}_{t-1}^{(i)}, \dots, \mathbf{x}_1^{(i)} \right], \\
&= \iota^{(i)'} \mathbf{0}.
\end{aligned}$$

(ii) Par définition (équation (2.1.1)), on a

$$V_t = \sigma^2 \frac{\prod_{i=1}^N C_t^{(i)}}{\prod_{i=1}^N \mathbb{E} \left[C_t^{(i)} \right]}.$$

Puisque $\mathbb{E} \left[C_t^{(i)} \right] = \boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)}$, on a

$$\begin{aligned}
V_t &= \frac{\sigma^2}{\prod_{i=1}^N \left[\boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)} \right]} \prod_{i=1}^N C_t^{(i)}, \\
&= \frac{\sigma^2}{\prod_{i=1}^N \left[\boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)} \right]} \exp \left(\log \left(\prod_{i=1}^N C_t^{(i)} \right) \right), \\
&= \frac{\sigma^2}{\prod_{i=1}^N \left[\boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)} \right]} \exp \left(\sum_{i=1}^N \log \left(C_t^{(i)} \right) \right), \\
&= \frac{\sigma^2}{\prod_{i=1}^N \left[\boldsymbol{\pi}^{(i)'} \boldsymbol{\nu}^{(i)} \right]} \exp \left(\sum_{i=1}^N h_t^{(i)} \right).
\end{aligned}$$

□

A.3. Preuve de la proposition 2.1.2

Proposition (Propriétés de la matrice de transition \mathbf{P}). *La matrice de transition $\mathbf{P} = \otimes_{i=1}^N (\mathbf{P}^{(i)})$ de V satisfait les propriétés suivantes :*

- (i) *La matrice de transition $\mathbf{P}^{(i)}$ de $C^{(i)}$ a une unique valeur propre non unitaire égale à $\phi^{(i)}$. La multiplicité algébrique de cette valeur propre est $K - 1$.*
- (ii) *La matrice de transition \mathbf{P} de V a 2^N valeurs propres données par :*

$$\left\{ \prod_{i \in I} \phi^{(i)} : I \in A \right\},$$

où A désigne l'ensemble des parties de $\{1, 2, \dots, N\}$. La multiplicité algébrique de la valeur propre $\prod_{i \in I} \phi^{(i)}$, $I \in A$ est $(K - 1)^{|I|}$.

- (iii) *La matrice de transition en k pas de temps \mathbf{P}^k de V admet la représentation suivante :*

$$\mathbf{P}^k = \otimes_{i=1}^N \left\{ \left(\phi^{(i)} \right)^k \mathbf{I}_K + \left[1 - \left(\phi^{(i)} \right)^k \right] \mathbf{1}_K \boldsymbol{\pi}^{(i)'} \right\},$$

et converge vers $\mathbf{\Pi} = \mathbf{1}_{K^N} \boldsymbol{\pi}'$ lorsque $k \rightarrow \infty$ avec un taux de convergence asymptotique donné par :

$$\mathbf{P}^k - \mathbf{\Pi} = \mathcal{O}(k^{K-2} a^k), \quad (\text{A.3.1})$$

et un taux de convergence non-asymptotique donné par :

$$\|\mathbf{P}^k - \mathbf{\Pi}\|_{\infty} \leq 2 - 2(1 - a^k)^N, \quad (\text{A.3.2})$$

où $a = \max_{i=1, \dots, N} \phi^{(i)}$ et $\|\cdot\|_{\infty}$ désigne la norme maximum des sommes des valeurs absolues des lignes.

DÉMONSTRATION. Soit $\mathbf{P}^{(i)}$ défini par l'équation (2.1.2) c'est-à-dire

$$\mathbf{P}^{(i)} = \phi^{(i)} \mathbf{I}_K + (1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'},$$

où \mathbf{I}_K est la matrice identité de dimension $K \times K$, $\mathbf{1}_K$ est le vecteur $K \times 1$ composé uniquement de 1, $\phi^{(i)} \in [0, 1)$ est un paramètre, et $\boldsymbol{\pi}^{(i)}$ est un vecteur paramètre de probabilités.

- (i) Puisque $\mathbf{P}^{(i)}$ est une matrice de transition, on a $\mathbf{P}^{(i)} \mathbf{1}_K = \mathbf{1}_K$. Donc, le vecteur $\mathbf{1}_K$ est un vecteur propre de $\mathbf{P}^{(i)}$ associé à la valeur 1. Nous allons montrer que $\phi^{(i)}$ est une valeur propre de $\mathbf{P}^{(i)}$ et que l'espace vectoriel $E_{\mathbf{P}^{(i)}}(\phi^{(i)})$ associé à $\phi^{(i)}$ est de dimension $K - 1$. Soit $\mathbf{x} = (x_1 \ x_2 \ \dots \ x_K)'$ un vecteur de l'espace vectoriel $E_{\mathbf{P}^{(i)}}(\phi^{(i)})$ c'est-à-dire vérifiant $(\mathbf{P}^{(i)} - \phi^{(i)} \mathbf{I}_K) \mathbf{x} = \mathbf{0}$. On a,

$$\begin{aligned} (\mathbf{P}^{(i)} - \phi^{(i)} \mathbf{I}_K) \mathbf{x} = \mathbf{0} &\iff (\phi^{(i)} \mathbf{I}_K + (1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'} - \phi^{(i)} \mathbf{I}_K) \mathbf{x} = \mathbf{0} \\ &\iff (1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'} \mathbf{x} = \mathbf{0} \\ &\iff \mathbf{1}_K \boldsymbol{\pi}^{(i)'} \mathbf{x} = \mathbf{0}. \end{aligned}$$

Puisque tous les éléments du vecteur à gauche sont identiques, nous pouvons simplifier le problème à l'équation

$$\sum_{j=1}^K \pi_j^{(i)} x_j = 0,$$

et donc les composantes de \mathbf{x} vérifient

$$x_K = - \sum_{j=1}^{K-1} \frac{\pi_j^{(i)}}{\pi_K^{(i)}} x_j.$$

Par suite, le vecteur \mathbf{x} vérifie la suite de relations

$$\begin{aligned} \mathbf{x} &= \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{K-1} \\ x_K \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{K-1} \\ - \sum_{j=1}^{K-1} \frac{\pi_j^{(i)}}{\pi_K^{(i)}} x_j \end{pmatrix} \\ &= x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{\pi_1^{(i)}}{\pi_K^{(i)}} \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -\frac{\pi_2^{(i)}}{\pi_K^{(i)}} \end{pmatrix} + \cdots + x_{K-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{\pi_{K-1}^{(i)}}{\pi_K^{(i)}} \end{pmatrix} = \sum_{j=1}^{K-1} x_j \mathbf{u}_j^{(i)}, \end{aligned}$$

où $\mathbf{u}_j^{(i)}$ est le vecteur ayant 1 à la j -ième composante, $-\frac{\pi_j^{(i)}}{\pi_K^{(i)}}$ à la K -ième composante et 0 partout ailleurs. Puisque les vecteurs $\mathbf{u}_j^{(i)}$, définis par

$$\mathbf{u}_1^{(i)} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ -\frac{\pi_1^{(i)}}{\pi_K^{(i)}} \end{pmatrix}, \mathbf{u}_2^{(i)} = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -\frac{\pi_2^{(i)}}{\pi_K^{(i)}} \end{pmatrix}, \dots, \mathbf{u}_{K-1}^{(i)} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -\frac{\pi_{K-1}^{(i)}}{\pi_K^{(i)}} \end{pmatrix}$$

sont linéairement indépendants, la famille $(\mathbf{u}_1^{(i)}, \dots, \mathbf{u}_{K-1}^{(i)})$ forme une base de l'espace vectoriel $E_{\mathbf{P}^{(i)}}(\phi^{(i)})$ et donc cet espace a pour dimension $K - 1$. Cet espace vectoriel n'étant pas égal à l'espace vectoriel trivial $\{\mathbf{0}\}$, on conclut que $\phi^{(i)}$ est une valeur propre de $\mathbf{P}^{(i)}$ et que la multiplicité géométrique de cette valeur propre est $K - 1$.

La dimension de la matrice $\mathbf{P}^{(i)}$ est K . Le valeur 1 est une valeur propre de $\mathbf{P}^{(i)}$ associée au vecteur propre $\mathbf{1}_K$. La valeur $\phi^{(i)}$ est également une valeur propre de $\mathbf{P}^{(i)}$ dont la multiplicité géométrique est $K - 1$. Puisque la multiplicité algébrique est supérieure ou égale à la multiplicité géométrique

alors la multiplicité algébrique de $\phi^{(i)}$ est aussi égale $K - 1$. En conclusion, la matrice de transition $\mathbf{P}^{(i)}$ de $C^{(i)}$ a une unique valeur propre non unitaire égale à $\phi^{(i)}$ et la multiplicité algébrique de cette valeur propre est $K - 1$.

(ii) Avant de démontrer la proposition, nous énonçons le résultat suivant :

Lemme 1. *Soit \mathbf{x} et \mathbf{y} , deux vecteurs propres des matrices \mathbf{A} et \mathbf{B} respectivement tels que $\mathbf{A}\mathbf{x} = \lambda_1\mathbf{x}$ et $\mathbf{B}\mathbf{y} = \lambda_2\mathbf{y}$. Le vecteur $\mathbf{x} \otimes \mathbf{y}$ est alors un vecteur propre de la matrice $\mathbf{A} \otimes \mathbf{B}$ associé à la valeur propre $\lambda_1\lambda_2$. En effet,*

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{x} \otimes \mathbf{y}) = (\mathbf{A}\mathbf{x}) \otimes (\mathbf{B}\mathbf{y}) = (\lambda_1\mathbf{x}) \otimes (\lambda_2\mathbf{y}) = \lambda_1\lambda_2(\mathbf{x} \otimes \mathbf{y}).$$

D'après (i), les vecteurs $\mathbf{u}_j^{(i)}$, $j = 1, 2, \dots, K - 1$ et le vecteur $\mathbf{u}_K^{(i)} = \mathbf{1}_K$ sont des vecteurs propres de $\mathbf{P}^{(i)}$ associés respectivement aux valeurs propres $\lambda_1^{(i)} = \lambda_2^{(i)} = \dots = \lambda_{K-1}^{(i)} = \phi^{(i)}$ et $\lambda_K^{(i)} = 1$. Ainsi, d'après le lemme 1, les vecteurs propres de $\mathbf{P} = \otimes_{i=1}^N (\mathbf{P}^{(i)})$ sont les vecteurs

$$\otimes_{i=1}^N (\mathbf{u}_{j_i}^{(i)}) = \mathbf{u}_{j_1}^{(1)} \otimes \mathbf{u}_{j_2}^{(2)} \otimes \dots \otimes \mathbf{u}_{j_N}^{(N)}$$

associés aux valeurs propres $\prod_{i=1}^N \lambda_{j_i}^{(i)}$ où pour tout $i = 1, 2, \dots, N$, $j_i \in \{1, 2, \dots, K\}$. L'ensemble des valeurs propres de \mathbf{P} est

$$\left\{ \prod_{i \in I} \phi^{(i)} : I \in A \right\},$$

où A désigne l'ensemble des parties de $\{1, 2, \dots, N\}$. En effet, la valeur propre $\prod_{i \in I} \phi^{(i)}$, $I \in A$ est obtenue si $j_i \in \{1, 2, \dots, K - 1\}$ pour tout $i \in I$ et $j_i = K$ pour tout $i \notin I$. Puisque la multiplicité de λ_{j_i} est $K - 1$ si $j_i \neq K$ et 1 si $j_i = K$, la multiplicité algébrique de la valeur propre $\prod_{i \in I} \phi^{(i)}$, $I \in A$ est $(K - 1)^{|I|}$.

(iii) Pour tout $i = 1, 2, \dots, N$ et $k = 1, 2, \dots$, on a :

$$\begin{aligned} [\mathbf{P}^{(i)}]^k &= \left[\phi^{(i)} \mathbf{I}_K + (1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'} \right]^k, \\ &= \sum_{j=0}^k \binom{k}{j} \left(\phi^{(i)} \mathbf{I}_K \right)^j \left((1 - \phi^{(i)}) \mathbf{1}_K \boldsymbol{\pi}^{(i)'} \right)^{k-j}, \\ &= \sum_{j=0}^k \binom{k}{j} \left(\phi^{(i)} \right)^j (\mathbf{I}_K)^j (1 - \phi^{(i)})^{k-j} \left(\mathbf{1}_K \boldsymbol{\pi}^{(i)'} \right)^{k-j}. \end{aligned}$$

Puisque les matrices \mathbf{I}_K et $\mathbf{1}_K \boldsymbol{\pi}^{(i) \prime}$ sont idempotentes c'est-à-dire $\mathbf{I}_K \mathbf{I}_K = \mathbf{I}_K$ et $\mathbf{1}_K \boldsymbol{\pi}^{(i) \prime} \mathbf{1}_K \boldsymbol{\pi}^{(i) \prime} = \mathbf{1}_K \boldsymbol{\pi}^{(i) \prime}$, on a :

$$\begin{aligned} [\mathbf{P}^{(i)}]^k &= (\phi^{(i)})^k \mathbf{I}_K + \sum_{j=0}^{k-1} \binom{k}{j} (\phi^{(i)})^j (1 - \phi^{(i)})^{k-j} (\mathbf{I}_K) (\mathbf{1}_K \boldsymbol{\pi}^{(i) \prime}), \\ &= (\phi^{(i)})^k \mathbf{I}_K + (\mathbf{1}_K \boldsymbol{\pi}^{(i) \prime}) \sum_{j=0}^{k-1} \binom{k}{j} (\phi^{(i)})^j (1 - \phi^{(i)})^{k-j}, \\ &= (\phi^{(i)})^k \mathbf{I}_K + \left[1 - (\phi^{(i)})^k\right] \mathbf{1}_K \boldsymbol{\pi}^{(i) \prime}. \end{aligned} \quad (\text{A.3.3})$$

On conclut donc que

$$\mathbf{P}^k = \otimes_{i=1}^N \left\{ (\phi^{(i)})^k \mathbf{I}_K + \left[1 - (\phi^{(i)})^k\right] \mathbf{1}_K \boldsymbol{\pi}^{(i) \prime} \right\}. \quad (\text{A.3.4})$$

De plus, lorsque k tend vers l'infini, $[\mathbf{P}^{(i)}]^k$ tend vers $\mathbf{1}_K \boldsymbol{\pi}^{(i) \prime}$ puisque $\phi^{(i)} \in [0, 1)$. Ainsi, \mathbf{P}^k converge vers $\otimes_{i=1}^N \left\{ \mathbf{1}_K \boldsymbol{\pi}^{(i) \prime} \right\} = (\otimes_{i=1}^N \mathbf{1}_K) (\otimes_{i=1}^N \boldsymbol{\pi}^{(i) \prime}) = \mathbf{1}_{K^N} \boldsymbol{\pi}' = \boldsymbol{\Pi}$.

Pour obtenir le taux de convergence non-asymptotique, remarquons à travers l'équation (A.3.4) que les termes matriciels du développement de \mathbf{P}^k sont les produits de Kronecker des combinaisons de N éléments de l'ensemble $I = \left\{ \mathbf{I}_K, \mathbf{1}_K \boldsymbol{\pi}^{(i) \prime}, i = 1, 2, \dots, N \right\}$ où seul \mathbf{I}_K peut être répété. Le dernier terme matriciel de ce développement est $\boldsymbol{\Pi}$ avec comme coefficient réel $\prod_{i=1}^N \left(1 - (\phi^{(i)})^k\right)$. Notons que tous les coefficients réels de ce développement sont positifs. Ainsi les termes matriciels de $\mathbf{P}^k - \boldsymbol{\Pi}$ sont les mêmes que ceux de \mathbf{P}^k , avec les mêmes coefficients réels, sauf le dernier terme $\boldsymbol{\Pi}$ qui aura pour coefficient réel $\left[\prod_{i=1}^N \left(1 - (\phi^{(i)})^k\right) - 1 \right]$ (négatif).

Pour être plus clair, désignons par $\mathbf{A}_l, l = 1, 2, \dots, 2^N$ les termes matriciels du développement de $\mathbf{P}^k - \boldsymbol{\Pi}$ et par α_l leurs coefficients réels respectifs. On a :

$$\mathbf{P}^k - \boldsymbol{\Pi} = \sum_{l=1}^{2^N} \alpha_l \mathbf{A}_l = \sum_{l=1}^{2^N-1} \alpha_l \mathbf{A}_l + \left[\prod_{i=1}^N \left(1 - (\phi^{(i)})^k\right) - 1 \right] \boldsymbol{\Pi}.$$

En utilisant l'inégalité triangulaire des normes, on a :

$$\begin{aligned} \|\mathbf{P}^k - \boldsymbol{\Pi}\|_\infty &\leq \sum_{l=1}^{2^N-1} \|\alpha_l \mathbf{A}_l\|_\infty + \left\| \left[\prod_{i=1}^N \left(1 - (\phi^{(i)})^k\right) - 1 \right] \boldsymbol{\Pi} \right\|_\infty, \\ &= \sum_{l=1}^{2^N-1} |\alpha_l| \cdot \|\mathbf{A}_l\|_\infty + \left| \prod_{i=1}^N \left(1 - (\phi^{(i)})^k\right) - 1 \right| \cdot \|\boldsymbol{\Pi}\|_\infty, \\ &= \sum_{l=1}^{2^N-1} \alpha_l \cdot \|\mathbf{A}_l\|_\infty + \left[1 - \prod_{i=1}^N \left(1 - (\phi^{(i)})^k\right) \right] \|\boldsymbol{\Pi}\|_\infty, \\ &= \sum_{l=1}^{2^N-1} \alpha_l + \left[1 - \prod_{i=1}^N \left(1 - (\phi^{(i)})^k\right) \right], \end{aligned}$$

puisque $\|\mathbf{A}_l\|_\infty = 1$. D'ailleurs, $\|\mathbf{A}\|_\infty = 1$ pour toute matrice $\mathbf{A} \in I$ et $\|\mathbf{A} \otimes \mathbf{B}\|_\infty = \|\mathbf{A}\|_\infty \|\mathbf{B}\|_\infty$ pour toutes matrices \mathbf{A} et \mathbf{B} . On a ensuite :

$$\begin{aligned}
\|\mathbf{P}^k - \mathbf{\Pi}\|_\infty &\leq \sum_{l=1}^{2^N-1} \alpha_l + \prod_{i=1}^N \left(1 - \left(\phi^{(i)}\right)^k\right) + \left[1 - 2 \prod_{i=1}^N \left(1 - \left(\phi^{(i)}\right)^k\right)\right], \\
&= 1 + \left[1 - 2 \prod_{i=1}^N \left(1 - \left(\phi^{(i)}\right)^k\right)\right], \\
&= 2 - 2 \prod_{i=1}^N \left(1 - \left(\phi^{(i)}\right)^k\right), \\
&\leq 2 - 2 \prod_{i=1}^N \left(1 - \left(\max_{i=1,2,\dots,N} \phi^{(i)}\right)^k\right), \\
&= 2 - 2(1 - a^k)^N.
\end{aligned} \tag{A.3.5}$$

L'équation (A.3.5) est due au fait que $\sum_{l=1}^{2^N-1} \alpha_l + \prod_{i=1}^N \left(1 - \left(\phi^{(i)}\right)^k\right) = 1$. En effet, il s'agit de la somme des coefficients réels des termes matriciels du développement de \mathbf{P}^k qu'on obtient à partir de (A.3.4). Ainsi,

$$\sum_{l=1}^{2^N-1} \alpha_l + \prod_{i=1}^N \left(1 - \left(\phi^{(i)}\right)^k\right) = \prod_{i=1}^N \left\{ \left(\phi^{(i)}\right)^k + \left[1 - \left(\phi^{(i)}\right)^k\right] \right\} = \prod_{i=1}^N \{1\} = 1.$$

Le taux de convergence asymptotique est une conséquence directe du théorème 1.2 de [Seneta \(2006\)](#) suivant :

Théorème. *Soit \mathbf{P} une matrice primitive, c'est-à-dire qu'il existe un entier positif l tel que tous les coefficients de \mathbf{P}^l sont strictement positifs. Si la deuxième plus grande valeur propre λ_2 de \mathbf{P} est non nulle ($\lambda_2 \neq 0$) alors, lorsque $k \rightarrow \infty$,*

$$\mathbf{P}^k = \lambda_1 \mathbf{u} \mathbf{v}' + \mathcal{O}(k^{m_2-1} |\lambda_2|^k),$$

où λ_1 est la plus grande valeur propre de \mathbf{P} , les vecteurs \mathbf{u} et \mathbf{v} sont les vecteurs propres gauche et droit de \mathbf{P} associés à λ_1 respectivement et m_2 est la multiplicité algébrique de la valeur propre λ_2 .

Dans le cadre du modèle MDSV, la plus grande valeur propre de \mathbf{P} est 1 et les vecteurs propres gauche et droit associés à cette valeur propre sont $\mathbf{1}_{K^N}$ et $\boldsymbol{\pi}$ puisque $\mathbf{P} \mathbf{1}_{K^N} = \mathbf{1}_{K^N}$ et $\boldsymbol{\pi}' \mathbf{P} = \boldsymbol{\pi}'$. La seconde plus grande valeur propre de \mathbf{P} est $a = \max_{i=1,\dots,N} \phi^{(i)}$ avec pour multiplicité algébrique $m_2 = K - 1$. On obtient donc

$$\mathbf{P}^k = \mathbf{1}_{K^N} \boldsymbol{\pi}' + \mathcal{O}(k^{K-2} a^k)$$

lorsque $k \rightarrow \infty$.

□

A.4. Preuve de la proposition 2.2.1

Proposition (Structure d'autocovariance du processus MDSV). *Soit φ un nombre réel et les deux quantités suivantes :*

$$\psi_\varphi = \frac{\omega(2 - \nu_0)^{2\varphi} + (1 - \omega)\nu_0^{2\varphi}}{[\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi]^2}, \quad \text{et} \quad \tau_\varphi = \frac{\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi}{[\omega(2 - \nu_0) + (1 - \omega)\nu_0]^\varphi}.$$

Pour $k \in \mathbb{N}$ et $k \geq 1$, on a :

(i)

$$\mathbb{E}[V_t^\varphi] = \sigma^{2\varphi} \tau_\varphi^{N(K-1)}. \quad (\text{A.4.1})$$

(ii)

$$\text{Var}(V_t^\varphi) = \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \left(\psi_\varphi^{N(K-1)} - 1 \right). \quad (\text{A.4.2})$$

(iii)

$$\text{Cov}(V_t^\varphi, V_{t+k}^\varphi) = \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \left(\prod_{i=1}^N \left(1 + (\psi_\varphi^{K-1} - 1) \phi^{(i)k} \right) - 1 \right). \quad (\text{A.4.3})$$

(iv)

$$\text{Corr}(V_t^\varphi, V_{t+k}^\varphi) = \frac{\prod_{i=1}^N \left(1 + (\psi_\varphi^{K-1} - 1) \phi^{(i)k} \right) - 1}{\psi_\varphi^{N(K-1)} - 1}. \quad (\text{A.4.4})$$

En particulier pour $\varphi = 1$, on a :

$$\psi_1 = \frac{\nu_0^2 + 4\omega(1 - \nu_0)}{(\nu_0 + 2\omega(1 - \nu_0))^2}, \quad \text{et} \quad \tau_1 = 1.$$

DÉMONSTRATION. Pour $i = 1, 2, \dots, N$, on désigne par $\nu^{(i)}$ et $\pi^{(i)}$ les vecteurs de volatilités et de probabilités stationnaires des chaînes de Markov $C^{(i)}$ définis par les équations (2.2.4) et (2.2.5), c'est-à-dire

$$\nu^{(i)} = \nu^{(1)} = \left(\nu_1^{(1)} \quad \nu_2^{(1)} \quad \dots \quad \nu_K^{(1)} \right)', \quad \nu_j^{(1)} = \nu_0 \left(\frac{2 - \nu_0}{\nu_0} \right)^{j-1}, \quad j = 1, \dots, K$$

et

$$\pi^{(i)} = \pi^{(1)} = \left(\pi_1^{(1)} \quad \pi_2^{(1)} \quad \dots \quad \pi_K^{(1)} \right), \quad \pi_j = \binom{K-1}{j-1} \omega^{j-1} (1 - \omega)^{K-j}, \quad j = 1, \dots, K,$$

où $\nu_0, \omega \in (0, 1)$. Désignons par \mathbf{d} le vecteur

$$\mathbf{d} = \left(\left(\nu_1^{(1)} \right)^\varphi \quad \left(\nu_2^{(1)} \right)^\varphi \quad \dots \quad \left(\nu_K^{(1)} \right)^\varphi \right)'$$

et par $\mathbf{\Upsilon}_D$, la matrice diagonale contenant \mathbf{d} c'est-à-dire vérifiant $\mathbf{\Upsilon}_D \mathbf{1}_K = \mathbf{d}$. Démontrons d'abord que

$$\frac{\pi^{(1)'} \mathbf{\Upsilon}_D \mathbf{d}}{(\pi^{(1)'} \mathbf{d})^2} = \psi_\varphi^{K-1}$$

et que

$$\frac{\pi^{(1)'} \mathbf{d}}{(\pi^{(1)'} \nu^{(1)})^\varphi} = \tau_\varphi^{K-1}.$$

Premièrement, on a

$$\begin{aligned}
\boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d} &= \sum_{j=1}^K \pi_j^{(1)} \left(\nu_j^{(1)} \right)^{2\varphi}, \\
&= \sum_{j=1}^K \binom{K-1}{j-1} \omega^{j-1} (1-\omega)^{K-j} \left(\nu_0 \left(\frac{2-\nu_0}{\nu_0} \right)^{j-1} \right)^{2\varphi}, \\
&= \nu_0^{2\varphi} \sum_{j=1}^K \binom{K-1}{j-1} (1-\omega)^{K-j} \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^{2\varphi} \right)^{j-1}, \\
&= \nu_0^{2\varphi} \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^{2\varphi} + 1 - \omega \right)^{K-1}.
\end{aligned}$$

Puis,

$$\begin{aligned}
\boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)} &= \sum_{j=1}^K \pi_j^{(1)} \nu_j^{(1)}, \\
&= \sum_{j=1}^K \binom{K-1}{j-1} \omega^{j-1} (1-\omega)^{K-j} \left(\nu_0 \left(\frac{2-\nu_0}{\nu_0} \right)^{j-1} \right), \\
&= \nu_0 \sum_{j=1}^K \binom{K-1}{j-1} (1-\omega)^{K-j} \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right) \right)^{j-1}, \\
&= \nu_0 \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right) + 1 - \omega \right)^{K-1}.
\end{aligned}$$

Ensuite

$$\begin{aligned}
\boldsymbol{\pi}^{(1)'} \mathbf{d} &= \sum_{j=1}^K \pi_j^{(1)} \left(\nu_j^{(1)} \right)^\varphi, \\
&= \sum_{j=1}^K \binom{K-1}{j-1} \omega^{j-1} (1-\omega)^{K-j} \left(\nu_0 \left(\frac{2-\nu_0}{\nu_0} \right)^{j-1} \right)^\varphi, \\
&= \nu_0^\varphi \sum_{j=1}^K \binom{K-1}{j-1} (1-\omega)^{K-j} \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^\varphi \right)^{j-1}, \\
&= \nu_0^\varphi \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^\varphi + 1 - \omega \right)^{K-1}.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\frac{\boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d}}{\left(\boldsymbol{\pi}^{(1)'} \mathbf{d} \right)^2} &= \frac{\nu_0^{2\varphi} \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^{2\varphi} + 1 - \omega \right)^{K-1}}{\nu_0^{2\varphi} \left[\left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^\varphi + 1 - \omega \right)^2 \right]^{K-1}}, \\
&= \left[\frac{\left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^{2\varphi} + 1 - \omega \right)}{\left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^\varphi + 1 - \omega \right)^2} \right]^{K-1}.
\end{aligned}$$

En multipliant le numérateur et le dénominateur par $(\nu_0^{2\varphi})^{K-1}$, on a

$$\frac{\boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d}}{(\boldsymbol{\pi}^{(1)'} \mathbf{d})^2} = \left[\frac{(\omega(2-\nu_0)^{2\varphi} + (1-\omega)\nu_0^{2\varphi})}{(\omega(2-\nu_0)^\varphi + (1-\omega)\nu_0^\varphi)^2} \right]^{K-1} = \psi_\varphi^{K-1}.$$

Aussi,

$$\begin{aligned} \frac{\boldsymbol{\pi}^{(1)'} \mathbf{d}}{(\boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)})^\varphi} &= \frac{\nu_0^\varphi \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^\varphi + 1 - \omega \right)^{K-1}}{\left[\nu_0 \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right) + 1 - \omega \right)^{K-1} \right]^\varphi}, \\ &= \frac{\nu_0^\varphi \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^\varphi + 1 - \omega \right)^{K-1}}{\nu_0^\varphi \left[\left(\omega \left(\frac{2-\nu_0}{\nu_0} \right) + 1 - \omega \right)^\varphi \right]^{K-1}}, \\ &= \left[\frac{\omega \left(\frac{2-\nu_0}{\nu_0} \right)^\varphi + 1 - \omega}{\left(\omega \left(\frac{2-\nu_0}{\nu_0} \right) + 1 - \omega \right)^\varphi} \right]^{K-1}. \end{aligned}$$

En multipliant le numérateur et le dénominateur par $(\nu_0^\varphi)^{K-1}$, on a

$$\frac{\boldsymbol{\pi}^{(1)'} \mathbf{d}}{(\boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)})^\varphi} = \left(\frac{\omega(2-\nu_0)^\varphi + (1-\omega)\nu_0^\varphi}{[\omega(2-\nu_0) + (1-\omega)\nu_0]^\varphi} \right)^{K-1} = \tau_\varphi^{K-1}.$$

Soit N chaînes de Markov homogènes et discrètes $\{D_t^{(1)}\}, \{D_t^{(2)}\}, \dots, \{D_t^{(N)}\}$, de matrices de transition respectives $\mathbf{P}^{(1)}, \mathbf{P}^{(2)}, \dots, \mathbf{P}^{(N)}$ et de même distribution stationnaire $\boldsymbol{\pi}^{(1)}$ telles que l'ensemble des valeurs prises par $\{D_t^{(i)}\}$, $i = 1, 2, \dots, N$, forment le vecteur \mathbf{d} . On peut constater que pour tout i , les variables $D_t^{(i)}$ et $(C_t^{(i)})^\varphi$ sont égales en loi. D'après la théorie des chaînes de Markov discrètes ([Hamilton, 1994](#)), on a :

$$\begin{aligned} \mathbb{E} \left[D_t^{(i)} \right] &= \boldsymbol{\pi}^{(1)'} \mathbf{d}, \\ \mathbb{E} \left[\left(D_t^{(i)} \right)^2 \right] &= \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d}. \end{aligned}$$

Pour $k = 1, 2, \dots$,

$$\mathbb{E} \left[D_t^{(i)} D_{t+k}^{(i)} \right] = \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{P}^{(i)k} \mathbf{d}.$$

où $\boldsymbol{\Upsilon}_D$ est la matrice diagonale contenant \mathbf{d} c'est-à-dire vérifiant $\boldsymbol{\Upsilon}_D \mathbf{1}_K = \mathbf{d}$. On a :

$$\begin{aligned}
\mathbb{E} \left[D_t^{(i)} D_{t+k}^{(i)} \right] &= \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{P}^{(i)k} \mathbf{d}, \\
&= \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \left\{ \left(\phi^{(i)} \right)^k \mathbf{I}_K + \left[1 - \left(\phi^{(i)} \right)^k \right] \mathbf{1}_K \boldsymbol{\pi}^{(1)'} \right\} \mathbf{d}, \\
&= \left(\phi^{(i)} \right)^k \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d} + \left[1 - \left(\phi^{(i)} \right)^k \right] \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{1}_K \boldsymbol{\pi}^{(1)'} \mathbf{d}, \\
&= \left(\phi^{(i)} \right)^k \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d} + \left[1 - \left(\phi^{(i)} \right)^k \right] \left(\boldsymbol{\pi}^{(1)'} \mathbf{d} \right)^2.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\frac{\mathbb{E} \left[D_t^{(i)} D_{t+k}^{(i)} \right]}{\mathbb{E} \left[D_t^{(i)} \right] \mathbb{E} \left[D_{t+k}^{(i)} \right]} &= \frac{\left(\phi^{(i)} \right)^k \boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d} + \left[1 - \left(\phi^{(i)} \right)^k \right] \left(\boldsymbol{\pi}^{(1)'} \mathbf{d} \right)^2}{\left(\boldsymbol{\pi}^{(1)'} \mathbf{d} \right)^2}, \\
&= \left(\phi^{(i)} \right)^k \frac{\boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d}}{\left(\boldsymbol{\pi}^{(1)'} \mathbf{d} \right)^2} + \left[1 - \left(\phi^{(i)} \right)^k \right], \\
&= 1 + \left(\phi^{(i)} \right)^k \left(\psi_\varphi^{K-1} - 1 \right).
\end{aligned} \tag{A.4.5}$$

(i) Puisque $V_t = \frac{\sigma^2}{\prod_{i=1}^N \mathbb{E} \left[C_t^{(i)} \right]} \prod_{i=1}^N C_t^{(i)}$, on a

$$\begin{aligned}
\mathbb{E} \left[V_t^\varphi \right] &= \frac{\sigma^{2\varphi}}{\left(\prod_{i=1}^N \mathbb{E} \left[C_t^{(i)} \right] \right)^\varphi} \mathbb{E} \left[\prod_{i=1}^N \left(C_t^{(i)} \right)^\varphi \right], \\
&= \sigma^{2\varphi} \prod_{i=1}^N \left[\frac{\mathbb{E} \left[\left(C_t^{(i)} \right)^\varphi \right]}{\left(\mathbb{E} \left[C_t^{(i)} \right] \right)^\varphi} \right].
\end{aligned} \tag{A.4.6}$$

En remarquant que $\mathbb{E} \left[\left(C_t^{(i)} \right)^\varphi \right] = \mathbb{E} \left[D_t^{(i)} \right] = \boldsymbol{\pi}^{(1)'} \mathbf{d}$ et que $\mathbb{E} \left[C_t^{(i)} \right] = \boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)}$, on a

$$\begin{aligned}
\mathbb{E} \left[V_t^\varphi \right] &= \sigma^{2\varphi} \prod_{i=1}^N \left[\frac{\boldsymbol{\pi}^{(1)'} \mathbf{d}}{\left(\boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)} \right)^\varphi} \right], \\
&= \sigma^{2\varphi} \left[\frac{\boldsymbol{\pi}^{(1)'} \mathbf{d}}{\left(\boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)} \right)^\varphi} \right]^N, \\
&= \sigma^{2\varphi} \left[\tau_\varphi^{K-1} \right]^N, \\
&= \sigma^{2\varphi} \tau_\varphi^{N(K-1)}.
\end{aligned}$$

(ii) En remplaçant φ par 2φ dans l'équation (A.4.6), on a :

$$\begin{aligned}
\mathbb{E} \left[V_t^{2\varphi} \right] &= \sigma^{4\varphi} \prod_{i=1}^N \left[\frac{\mathbb{E} \left[\left(C_t^{(i)} \right)^{2\varphi} \right]}{\left(\mathbb{E} \left[C_t^{(i)} \right] \right)^{2\varphi}} \right], \\
&= \sigma^{4\varphi} \prod_{i=1}^N \left[\frac{\mathbb{E} \left[\left(D_t^{(i)} \right)^2 \right]}{\left(\mathbb{E} \left[C_t^{(i)} \right] \right)^{2\varphi}} \right], \\
&= \sigma^{4\varphi} \left[\frac{\boldsymbol{\pi}^{(1)'} \boldsymbol{\Upsilon}_D \mathbf{d}}{\left(\boldsymbol{\pi}^{(1)'} \boldsymbol{\nu}^{(1)} \right)^{2\varphi}} \right]^N, \\
&= \sigma^{4\varphi} \left[\frac{\nu_0^{2\varphi} \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right)^{2\varphi} + 1 - \omega \right)^{K-1}}{\left(\nu_0 \left(\omega \left(\frac{2-\nu_0}{\nu_0} \right) + 1 - \omega \right)^{K-1} \right)^{2\varphi}} \right]^N.
\end{aligned}$$

Après simplification, on obtient

$$\begin{aligned}
\mathbb{E} \left[V_t^{2\varphi} \right] &= \sigma^{4\varphi} \left[\frac{\omega (2 - \nu_0)^{2\varphi} + (1 - \omega) \nu_0^{2\varphi}}{(\omega (2 - \nu_0) + (1 - \omega) \nu_0)^{2\varphi}} \right]^{N(K-1)}, \\
&= \sigma^{4\varphi} \left[\frac{\omega (2 - \nu_0)^{2\varphi} + (1 - \omega) \nu_0^{2\varphi}}{(\omega (2 - \nu_0)^\varphi + (1 - \omega) \nu_0^\varphi)^2} \times \frac{(\omega (2 - \nu_0)^\varphi + (1 - \omega) \nu_0^\varphi)^2}{(\omega (2 - \nu_0) + (1 - \omega) \nu_0)^{2\varphi}} \right]^{N(K-1)}, \\
&= \sigma^{4\varphi} \psi_\varphi^{N(K-1)} \tau_\varphi^{2N(K-1)}. \tag{A.4.7}
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\text{Var} (V_t^\varphi) &= \mathbb{E} \left[V_t^{2\varphi} \right] - \left(\mathbb{E} [V_t^\varphi] \right)^2, \\
&= \sigma^{4\varphi} \psi_\varphi^{N(K-1)} \tau_\varphi^{2N(K-1)} - \left(\sigma^{2\varphi} \tau_\varphi^{N(K-1)} \right)^2, \\
&= \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \left[\psi_\varphi^{N(K-1)} - 1 \right].
\end{aligned}$$

(iii) On a :

$$\mathbb{E}[V_t^\varphi V_{t+k}^\varphi] = \frac{\sigma^{4\varphi}}{\left(\prod_{i=1}^N \mathbb{E} \left[C_t^{(i)} \right] \right)^{2\varphi}} \mathbb{E} \left[\prod_{i=1}^N \left(C_t^{(i)} \right)^\varphi \left(C_{t+k}^{(i)} \right)^\varphi \right].$$

Par indépendance des chaînes de Markov $C^{(i)}$, on a

$$\begin{aligned}
\mathbb{E}[V_t^\varphi V_{t+k}^\varphi] &= \sigma^{4\varphi} \frac{\left(\prod_{i=1}^N \mathbb{E} \left[\left(C_t^{(i)} \right)^\varphi \right]\right)^2}{\left(\prod_{i=1}^N \mathbb{E} \left[C_t^{(i)} \right]\right)^{2\varphi}} \prod_{i=1}^N \left(\frac{\mathbb{E} \left[\left(C_t^{(i)} \right)^\varphi \left(C_{t+k}^{(i)} \right)^\varphi \right]}{\mathbb{E} \left[\left(C_t^{(i)} \right)^\varphi \right] \mathbb{E} \left[\left(C_{t+k}^{(i)} \right)^\varphi \right]} \right), \\
&= \sigma^{4\varphi} \prod_{i=1}^N \left[\frac{\mathbb{E} \left[\left(C_t^{(i)} \right)^\varphi \right]}{\left(\mathbb{E} \left[C_t^{(i)} \right]\right)^\varphi} \right]^2 \prod_{i=1}^N \left(\frac{\mathbb{E} \left[D_t^{(i)} D_{t+k}^{(i)} \right]}{\mathbb{E} \left[D_t^{(i)} \right] \mathbb{E} \left[D_{t+k}^{(i)} \right]} \right), \\
&= \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \prod_{i=1}^N \left[1 + \left(\phi^{(i)} \right)^k \left(\psi_\varphi^{K-1} - 1 \right) \right].
\end{aligned}$$

La dernière équation est obtenue puisque (du calcul de $\mathbb{E}[V_t^\varphi]$ (A.4.6))

$$\frac{\mathbb{E} \left[\left(C_t^{(i)} \right)^\varphi \right]}{\left(\mathbb{E} \left[C_t^{(i)} \right]\right)^\varphi} = \tau_\varphi^{N(K-1)}$$

et (de l'équation (A.4.5))

$$\frac{\mathbb{E} \left[D_t^{(i)} D_{t+k}^{(i)} \right]}{\mathbb{E} \left[D_t^{(i)} \right] \mathbb{E} \left[D_{t+k}^{(i)} \right]} = 1 + \left(\phi^{(i)} \right)^k \left(\psi_\varphi^{K-1} - 1 \right).$$

Par la suite,

$$\begin{aligned}
\text{Cov} \left(V_t^\varphi, V_{t+k}^\varphi \right) &= \mathbb{E} \left[V_t^\varphi V_{t+k}^\varphi \right] - \mathbb{E} \left[V_t^\varphi \right] \mathbb{E} \left[V_{t+k}^\varphi \right], \\
&= \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \prod_{i=1}^N \left[1 + \left(\phi^{(i)} \right)^k \left(\psi_\varphi^{K-1} - 1 \right) \right] - \left(\sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \right), \\
&= \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} \left(\prod_{i=1}^N \left(1 + \left(\phi^{(i)} \right)^k \left(\psi_\varphi^{K-1} - 1 \right) \right) - 1 \right).
\end{aligned}$$

(iv) Le résultat s'obtient directement à partir de la formule de la fonction de corrélation.

Le cas particulier de $\varphi = 1$ donne

$$\begin{aligned}
\psi_\varphi &= \frac{\omega(2 - \nu_0)^{2\varphi} + (1 - \omega)\nu_0^{2\varphi}}{[\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi]^2}, \\
&= \frac{\omega(2 - \nu_0)^2 + \nu_0^2 - \nu_0^2\omega}{(\omega(2 - \nu_0) + \nu_0 - \nu_0\omega)^2}, \\
&= \frac{\nu_0^2 + \omega \left((2 - \nu_0)^2 - \nu_0^2 \right)}{(\nu_0 + 2\omega(1 - \nu_0))^2}, \\
&= \frac{\nu_0^2 + 4\omega(1 - \nu_0)}{(\nu_0 + 2\omega(1 - \nu_0))^2}.
\end{aligned}$$

et

$$\begin{aligned}\tau_\varphi &= \frac{\omega(2-\nu_0)^\varphi + (1-\omega)\nu_0^\varphi}{[\omega(2-\nu_0) + (1-\omega)\nu_0]^\varphi}, \\ &= \frac{\omega(2-\nu_0) + (1-\omega)\nu_0}{\omega(2-\nu_0) + (1-\omega)\nu_0}, \\ &= 1.\end{aligned}$$

□

A.5. Preuve de la proposition 2.2.2

Proposition (Distribution stationnaire et distribution du temps de séjour). *L'espace d'états ν de dimension K^N du processus V défini à l'équation (2.2.1) contient $N(K-1)+1$ valeurs distinctes, notées dans un ordre croissant par $x_1, x_2, \dots, x_{N(K-1)+1}$ où :*

$$x_k = \nu_0^{N-k+1}(2-\nu_0)^{k-1}, \quad k = 1, 2, \dots, N(K-1)+1.$$

(i) *Le nombre n_k de fois que la valeur x_k est répétée dans l'espace d'états ν est*

$$n_k = \sum_{m=0}^{\lfloor \frac{k-1}{K} \rfloor} (-1)^m \binom{N+k-mK-2}{N-1} \binom{N}{m},$$

où le symbole $\lfloor \cdot \rfloor$ représente la partie entière.

(ii) *La distribution stationnaire du processus V sur $x_1, x_2, \dots, x_{N(K-1)+1}$ présente une structure binomiale :*

$$\mathbb{P}[V_t = x_k] = \binom{N(K-1)}{k-1} \omega^{k-1} (1-\omega)^{N(K-1)+1-k}, \quad k = 1, 2, \dots, N(K-1)+1.$$

(iii) *La distribution du temps de séjour du processus V dans le sous-ensemble des états ayant pour valeur x_k est donnée par :*

$$\mathbb{P}(N_k = n) = \frac{\mathbf{1}'_{n_k} (\mathbf{I}_{n_k} - \mathbf{P}_{x_k}) \mathbf{P}_{x_k}^{n-1} (\mathbf{I}_{n_k} - \mathbf{P}_{x_k}) \mathbf{1}_{n_k}}{\mathbf{1}'_{n_k} (\mathbf{I}_{n_k} - \mathbf{P}_{x_k}) \mathbf{1}_{n_k}}, \quad n = 1, 2, \dots,$$

où N_k désigne la variable aléatoire qui compte le nombre de périodes successives pendant lesquelles la valeur prise par le processus V est x_k et \mathbf{P}_{x_k} représente la sous-matrice de \mathbf{P} qui contient les probabilités de transitions dans le sous-espace d'états de valeur x_k .

DÉMONSTRATION. Les valeurs prises par V_t sont les produits de N éléments (qui peuvent être choisis avec répétition) de l'ensemble $\{\nu_1, \nu_2, \dots, \nu_K\}$. Ainsi, ces valeurs peuvent être représentées par

$$\prod_{i=1}^N \nu_{j_i}, \quad j_1, j_2, \dots, j_N \in \{1, 2, \dots, K\}.$$

De plus, on a :

$$\prod_{i=1}^N \nu_{j_i} = \prod_{i=1}^N \left[\nu_0 \left(\frac{2-\nu_0}{\nu_0} \right)^{j_i-1} \right] = \nu_0^N \left(\frac{2-\nu_0}{\nu_0} \right)^{\sum_{i=1}^N j_i - N}.$$

Finalement, le nombre de valeurs différentes que peut prendre V_t est le nombre de valeurs possibles de la somme $\sum_{i=1}^N j_i$ si $j_1, j_2, \dots, j_N \in \{1, 2, \dots, K\}$. Les valeurs possibles de $\sum_{i=1}^N j_i$ sont $N, N+1, \dots, NK$. Ainsi, V_t peut prendre $NK - N + 1$ valeurs différentes qui sont

$$x_k = \nu_0^N \left(\frac{2 - \nu_0}{\nu_0} \right)^{k-1},$$

avec $k = \sum_{i=1}^N j_i - N + 1$, c'est-à-dire $k \in \{1, 2, \dots, N(K-1) + 1\}$.

(i) Pour une valeur fixe de k de l'ensemble $\{1, 2, \dots, N(K-1) + 1\}$, la détermination du nombre n_k de fois que la valeur x_k est répétée dans le vecteur ν peut donc se ramener à un problème de dénombrement du nombre de façons de résoudre l'équation

$$\sum_{i=1}^N j_i = k + N - 1,$$

avec $j_i \in \mathbb{N}$ et $1 \leq j_i \leq K$. Pour répondre à cette question, on fait un changement de variable en posant $y_i = j_i - 1$. L'équation devient ainsi

$$y_1 + y_2 + \dots + y_N = k - 1,$$

avec $y_i \in \mathbb{N}$ et $0 \leq y_i \leq K - 1$. Soit U l'ensemble des solutions entières de l'équation

$$y_1 + y_2 + \dots + y_N = k - 1,$$

où on exige seulement que $y_i \in \mathbb{N}$ pour chaque i . Pour chaque i , $1 \leq i \leq N$, définissons $U_i \subseteq U$ comme le sous-ensemble des solutions où $y_i \geq K$ (donc on exige seulement une extra condition). Donc $U - (U_1 \cup U_2 \cup \dots \cup U_N)$ est l'ensemble des solutions qui satisfont aucune des conditions $y_i \geq K$; c'est l'ensemble qui nous intéresse. Nous cherchons donc $|U| - |U_1 \cup U_2 \cup \dots \cup U_N|$. On sait que

$$|U| = \binom{N+k-2}{N-1}$$

et pour tout i ,

$$|U_i| = \binom{N+k-K-2}{N-1} \mathbf{1}_{\{k-1 \geq K\}}.$$

De plus, pour tout sous ensemble $1 \leq i_1 \leq i_2 \leq \dots \leq i_m \leq N$ on a

$$|U_{i_1} \cap \dots \cap U_{i_m}| = \binom{N+k-mK-2}{N-1} \mathbf{1}_{\{k-1 \geq mK\}}.$$

Donc en utilisant le principe d'inclusion exclusion, on obtient

$$n_k = |U| - |U_1 \cup U_2 \cup \dots \cup U_N| = \sum_{m=0}^{\lfloor \frac{k-1}{K} \rfloor} (-1)^m \binom{N+k-mK-2}{N-1} \binom{N}{m}.$$

(ii) La proposition 2.2.2 (ii) est démontrée par induction sur N . Pour cela, posons $V_t^{(N)} = \prod_{i=1}^N C_t^{(i)}$.

Début ($N = 1$) : On a $V_t^{(1)} = C_t^{(1)}$ et la distribution stationnaire de $C_t^{(1)}$ est binomiale :

$$\mathbb{P} \left[V_t^{(N)} = x_j \right] = \mathbb{P} \left[C_t^{(1)} = x_j \right] = \binom{K-1}{j-1} \omega^{j-1} (1-\omega)^{K-j}, \quad j = 1, \dots, K.$$

Étape d'induction : Supposons, pour N fixé, que la distribution stationnaire de $V_t^{(N)}$ soit binomiale :

$$\mathbb{P} \left[V_t^{(N)} = u_i \right] = \binom{N(K-1)}{i-1} \omega^{i-1} (1-\omega)^{N(K-1)+1-i}, \quad i = 1, \dots, N(K-1) + 1,$$

et

$$u_i = \nu_0^N \left(\frac{2-\nu_0}{\nu_0} \right)^{i-1}, \quad i = 1, \dots, N(K-1) + 1.$$

En supposant qu'on a une chaîne de Markov $C^{(N+1)}$ supplémentaire prenant ses valeurs dans $\{\nu_1, \nu_2, \dots, \nu_K\}$ où

$$\nu_j = \nu_0 \left(\frac{2-\nu_0}{\nu_0} \right)^{j-1}, \quad j = 1, \dots, K,$$

avec pour probabilité stationnaire $\{\pi_1, \pi_2, \dots, \pi_K\}$ où

$$\mathbb{P} \left[C_t^{(N+1)} = \nu_j \right] = \pi_j = \binom{K-1}{j-1} \omega^{j-1} (1-\omega)^{K-j}, \quad j = 1, \dots, K,$$

et que

$$V_t^{(N+1)} = \prod_{i=1}^{N+1} C_t^{(i)} = V_t^{(N)} C_t^{(N+1)},$$

déterminons la distribution stationnaire de $V_t^{(N+1)}$.

Les valeurs prises par $V_t^{(N+1)}$ sont les produits $u_i \nu_j$ où $i \in \{1, \dots, N(K-1) + 1\}$ et $j \in \{1, \dots, K\}$. On a :

$$u_i \nu_j = \nu_0^N \left(\frac{2-\nu_0}{\nu_0} \right)^{i-1} \nu_0 \left(\frac{2-\nu_0}{\nu_0} \right)^{j-1} = \nu_0^{N+1} \left(\frac{2-\nu_0}{\nu_0} \right)^{i+j-2}.$$

En posant $k = i + j - 1$, on a

$$x_k = u_i \nu_j = \nu_0^{N+1} \left(\frac{2-\nu_0}{\nu_0} \right)^{k-1},$$

avec $k \in \{1, 2, \dots, (N+1)(K-1) + 1\}$.

Pour i et j fixé, la probabilité que $V_t^{(N)} = u_i$ et $C_t^{(N+1)} = \nu_j$ est :

$$\begin{aligned} \mathbb{P} \left[V_t^{(N)} = u_i, C_t^{(i)} = \nu_j \right] &= \mathbb{P} \left[V_t^{(N)} = u_i \right] \mathbb{P} \left[C_t^{(i)} = \nu_j \right], \\ &= \binom{N(K-1)}{i-1} \omega^{i-1} (1-\omega)^{N(K-1)+1-i} \binom{K-1}{j-1} \omega^{j-1} (1-\omega)^{K-j}, \\ &= \binom{N(K-1)}{i-1} \binom{K-1}{j-1} \omega^{(i+j-1)-1} (1-\omega)^{N(K-1)+K-(i+j-1)}, \end{aligned}$$

où la deuxième égalité s'obtient par indépendance des chaînes de Markov.

Ainsi, pour tout $k \in \{1, 2, \dots, (N+1)(K-1)+1\}$, on a :

$$\begin{aligned}
\mathbb{P} \left[V_t^{(N+1)} = x_k \right] &= \sum_{i=1}^{N(K-1)+1} \sum_{j=1}^K \mathbf{1}_{\{k=i+j-1\}} \mathbb{P} \left[V_t^{(N)} = u_i, C_t^{(i)} = \nu_j \right], \\
&= \sum_{i=1}^{N(K-1)+1} \sum_{j=1}^K \mathbf{1}_{\{k=i+j-1\}} \binom{N(K-1)}{i-1} \binom{K-1}{j-1} \omega^{(i+j-1)-1} (1-\omega)^{N(K-1)+K-(i+j-1)}, \\
&= \sum_{i=1}^k \binom{N(K-1)}{i-1} \binom{K-1}{k-i} \omega^{k-1} (1-\omega)^{N(K-1)+K-k}, \\
&= \left[\omega^{k-1} (1-\omega)^{N(K-1)+K-k} \right] \sum_{i=1}^k \binom{N(K-1)}{i-1} \binom{K-1}{k-i}.
\end{aligned}$$

Pour la troisième égalité, $j-1 = k-i$ et la valeur la plus grande possible de i pour que $k = i+j-1$ est k . De plus, d'après l'identité de Vandermonde, pour tout entier m, n et r tels que $r \leq m$, on a

$$\binom{n+m}{r} = \sum_{i=0}^r \binom{n}{i} \binom{m}{r-i}.$$

Donc

$$\begin{aligned}
\sum_{i=1}^k \binom{N(K-1)}{i-1} \binom{K-1}{k-i} &= \sum_{i=1}^k \binom{N(K-1)}{i-1} \binom{K-1}{(k-1)-(i-1)}, \\
&= \sum_{i=0}^{k-1} \binom{N(K-1)}{i} \binom{K-1}{(k-1)-i}, \\
&= \binom{N(K-1)+K-1}{k-1}, \\
&= \binom{(N+1)(K-1)}{k-1}.
\end{aligned}$$

Il s'ensuit que,

$$\mathbb{P} \left[V_t^{(N+1)} = x_k \right] = \binom{(N+1)(K-1)}{k-1} \omega^{k-1} (1-\omega)^{(N+1)(K-1)+1-k}.$$

Conclusion : La distribution stationnaire de $V_t^{(N)}$ est binomiale c'est-à-dire

$$\mathbb{P} \left[V_t^{(N)} = x_k \right] = \binom{N(K-1)}{k-1} \omega^{k-1} (1-\omega)^{N(K-1)+1-k}, \quad k = 1, \dots, N(K-1)+1,$$

où

$$x_k = \nu_0^N \left(\frac{2-\nu_0}{\nu_0} \right)^{k-1}, \quad k = 1, \dots, N(K-1)+1.$$

(iii) Ce résultat découle du théorème 2.1.2. Il suffit de poser ν_s le sous ensemble de ν dont toutes les valeurs sont égales à x_k et remarquer que $\pi_s = \binom{N(K-1)}{k-1} \omega^{k-1} (1-\omega)^{N(K-1)+1-k} \mathbf{1}_K$. Le résultat s'obtient directement.

□

A.6. Preuve de la proposition 2.3.1

Proposition. (Structure d'autocovariance de $\{r_t^2\}$, $\{RV_t\}$ et $\{(r_t^2, RV_t)\}$) Soit les deux quantités suivantes :

$$\psi_\varphi = \frac{\omega(2 - \nu_0)^{2\varphi} + (1 - \omega)\nu_0^{2\varphi}}{[\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi]^2}, \quad \text{et} \quad \tau_\varphi = \frac{\omega(2 - \nu_0)^\varphi + (1 - \omega)\nu_0^\varphi}{[\omega(2 - \nu_0) + (1 - \omega)\nu_0]^\varphi}.$$

Pour $k \in \mathbb{N}$,

(i) si r_t^2 représente le carré des rendements financiers, où r_t est défini par l'équation (2.3.1), alors

$$\begin{aligned} \text{Cov}(r_t^2, r_{t+k}^2) &= \text{Cov}(V_t, V_{t+k}), \\ \text{Var}(r_t^2) &= \sigma^4 \left(\mathbb{E}[\epsilon_t^4] \psi_1^{N(K-1)} - 1 \right). \end{aligned}$$

Si le processus $\{\epsilon_t\}$ est distribué selon une loi normale centrée réduite alors $\mathbb{E}[\epsilon_t^4] = 3$.

(ii) si RV_t représente les variances réalisées, définies par l'équation (2.3.2), alors

$$\begin{aligned} \text{Cov}(RV_t, RV_{t+k}) &= \text{Cov}(V_t, V_{t+k}), \\ \text{Var}(RV_t) &= \sigma^4 \left(\mathbb{E}[\eta_t^2] \psi_1^{N(K-1)} - 1 \right). \end{aligned}$$

Si le processus $\{\eta_t\}$ est distribué selon une loi Gamma ($\text{Gamma}(\nu, 1/\nu)$) de moyenne 1 et de paramètre de forme ν alors $\mathbb{E}[\eta_t^2] = \frac{\nu + 1}{\nu}$.

(iii) si (r_t, RV_t) est défini par les équations (2.3.3) et (2.3.4), alors

$$\begin{aligned} \text{Cov}(r_t^2, r_{t+k}^2) &= \text{Cov}(V_t, V_{t+k}), \\ \text{Var}(r_t^2) &= \sigma^4 \left(\mathbb{E}[\epsilon_t^4] \psi_1^{N(K-1)} - 1 \right), \end{aligned}$$

et si $\delta_2 < \frac{1}{2}$, on a :

$$\begin{aligned} \text{Cov}(RV_t, RV_{t+k}) &= e^{2\xi} (\mathbb{E}[\eta_t])^2 \text{Cov}(V_t^\varphi, V_{t+k}^\varphi), \\ \text{Var}(RV_t) &= \sigma^{4\varphi} e^{2\xi} \tau_\varphi^{2N(K-1)} \left(\psi_\varphi^{N(K-1)} \mathbb{E}[\eta_t^2] - (\mathbb{E}[\eta_t])^2 \right), \\ \text{Cov}(RV_t, V_t) &= \sigma^{2(\varphi+1)} e^\xi \mathbb{E}[\eta_t] \left(\tau_{\varphi+1}^{N(K-1)} - \tau_\varphi^{N(K-1)} \right). \end{aligned}$$

Dans ces équations, $\eta_t = \exp(\delta_1 \epsilon_t + \delta_2 \epsilon_t^2 + \gamma \epsilon_t - \delta_2)$ et $\text{Cov}(V_t^\varphi, V_{t+k}^\varphi)$ est donné dans la proposition 2.2.1. De plus, si les processus $\{\varepsilon_t\}$ et $\{\epsilon_t\}$ sont normalement distribués alors

$$\mathbb{E}[\eta_t] = \frac{\exp\left(\frac{\delta_1^2}{2-4\delta_2} + \frac{\gamma^2}{2} - \delta_2\right)}{\sqrt{1-2\delta_2}} \quad \text{et} \quad \mathbb{E}[\eta_t^2] = \frac{\exp\left(\frac{2\delta_1^2}{1-4\delta_2} + 2\gamma^2 - 2\delta_2\right)}{\sqrt{1-4\delta_2}}.$$

DÉMONSTRATION. Soit $k \in \mathbb{N}$.

(i) Posons $y_t = r_t^2$. On a :

$$\begin{aligned} \mathbb{E}[y_t] &= \mathbb{E}[V_t \epsilon_t^2] = \mathbb{E}[V_t] \mathbb{E}[\epsilon_t^2] = \mathbb{E}[V_t], \\ \mathbb{E}[y_t y_{t+k}] &= \mathbb{E}[V_t \epsilon_t^2 V_{t+k} \epsilon_{t+k}^2] = \mathbb{E}[V_t V_{t+k}] \mathbb{E}[\epsilon_t^2] \mathbb{E}[\epsilon_{t+k}^2] = \mathbb{E}[V_t V_{t+k}]. \end{aligned}$$

et d'après (A.4.7) avec $\varphi = 1$, on a

$$\mathbb{E}[y_t^2] = \mathbb{E}[V_t^2 \epsilon_t^4] = \mathbb{E}[V_t^2] \mathbb{E}[\epsilon_t^4] = \sigma^4 \psi_1^{N(K-1)} \mathbb{E}[\epsilon_t^4].$$

On obtient donc

$$\begin{aligned} \mathbb{Cov}(y_t, y_{t+k}) &= \mathbb{E}[V_t V_{t+k}] - \mathbb{E}[V_t] \mathbb{E}[V_{t+k}] = \mathbb{Cov}(V_t, V_{t+k}), \\ \mathbb{Var}[y_t] &= \mathbb{E}[y_t^2] - (\mathbb{E}[y_t])^2 = \sigma^4 \left(\psi_1^{N(K-1)} \mathbb{E}[\epsilon_t^4] - 1 \right). \end{aligned}$$

(ii) La démonstration pour $y_t = RV_t$ est similaire à celle au point (i). Il suffit de remplacer η_t et η_{t+k} par ϵ_t^2 et ϵ_{t+k}^2 respectivement.

(iii) Comme précédemment, on a $\mathbb{E}[r_t^2] = \mathbb{E}[V_t]$ et $\mathbb{E}[r_t^2 r_{t+k}^2] = \mathbb{E}[V_t V_{t+k}]$. Donc

$$\begin{aligned} \mathbb{Cov}(r_t^2, r_{t+k}^2) &= \mathbb{Cov}(V_t, V_{t+k}), \\ \mathbb{Var}(r_t^2) &= \sigma^4 \left(\psi_1^{N(K-1)} \mathbb{E}[\epsilon_t^4] - 1 \right). \end{aligned}$$

L'équation (2.3.4) peut se réécrire

$$RV_t = e^\xi V_t^\varphi \eta_t,$$

où $\eta_t = \exp(\delta_1 \epsilon_t + \delta_2 \epsilon_t^2 + \gamma \epsilon_t - \delta_2)$. On a donc :

$$\begin{aligned} \mathbb{E}[RV_t] &= \mathbb{E}[e^\xi V_t^\varphi \eta_t] = e^\xi \mathbb{E}[V_t^\varphi] \mathbb{E}[\eta_t], \\ \mathbb{E}[RV_t V_t] &= \mathbb{E}[e^\xi V_t^\varphi \eta_t V_t] = e^\xi \mathbb{E}[V_t^{\varphi+1}] \mathbb{E}[\eta_t], \end{aligned}$$

et

$$\begin{aligned} \mathbb{E}[RV_t RV_{t+k}] &= \mathbb{E}[e^\xi V_t^\varphi \eta_t e^\xi V_{t+k}^\varphi \eta_{t+k}] = e^{2\xi} \mathbb{E}[V_t^\varphi V_{t+k}^\varphi] \mathbb{E}[\eta_t] \mathbb{E}[\eta_{t+k}], \\ &= e^{2\xi} \mathbb{E}[V_t^\varphi V_{t+k}^\varphi] (\mathbb{E}[\eta_t])^2. \end{aligned}$$

On obtient ainsi

$$\begin{aligned} \mathbb{Cov}(RV_t, RV_{t+k}) &= \mathbb{E}[RV_t RV_{t+k}] - (\mathbb{E}[RV_t])^2, \\ &= e^{2\xi} \mathbb{E}[V_t^\varphi V_{t+k}^\varphi] (\mathbb{E}[\eta_t])^2 - e^{2\xi} (\mathbb{E}[V_t^\varphi])^2 (\mathbb{E}[\eta_t])^2, \\ &= e^{2\xi} (\mathbb{E}[\eta_t])^2 \left[\mathbb{E}[V_t^\varphi V_{t+k}^\varphi] - (\mathbb{E}[V_t^\varphi])^2 \right], \\ &= e^{2\xi} (\mathbb{E}[\eta_t])^2 \mathbb{Cov}(V_t^\varphi, V_{t+k}^\varphi). \end{aligned}$$

De plus,

$$\begin{aligned}
\text{Cov}(RV_t, V_t) &= \mathbb{E}[RV_t V_t] - (\mathbb{E}[RV_t]) (\mathbb{E}[V_t]) , \\
&= e^\xi \mathbb{E} \left[V_t^{\varphi+1} \right] \mathbb{E}[\eta_t] - (e^\xi \mathbb{E}[V_t^\varphi] \mathbb{E}[\eta_t]) (\sigma^2) , \\
&= e^\xi \mathbb{E}[\eta_t] \left(\sigma^{2(\varphi+1)} \tau_{\varphi+1}^{N(K-1)} - \sigma^{2(\varphi+1)} \tau_\varphi^{N(K-1)} \right) , \\
&= \sigma^{2(\varphi+1)} e^\xi \mathbb{E}[\eta_t] \left(\tau_{\varphi+1}^{N(K-1)} - \tau_\varphi^{N(K-1)} \right) .
\end{aligned}$$

Pour la variance, on a

$$\mathbb{E}[RV_t^2] = \mathbb{E} \left[e^{2\xi} V_t^{2\varphi} \eta_t^2 \right] = e^{2\xi} \mathbb{E} \left[V_t^{2\varphi} \right] \mathbb{E}[\eta_t^2] ,$$

et

$$\begin{aligned}
\text{Var}(RV_t) &= \mathbb{E}[RV_t^2] - (\mathbb{E}[RV_t])^2 , \\
&= e^{2\xi} \mathbb{E} \left[V_t^{2\varphi} \right] \mathbb{E}[\eta_t^2] - e^{2\xi} (\mathbb{E}[V_t^\varphi])^2 (\mathbb{E}[\eta_t])^2 , \\
&= e^{2\xi} \left[\sigma^{4\varphi} \psi_\varphi^{N(K-1)} \tau_\varphi^{2N(K-1)} \mathbb{E}[\eta_t^2] - \sigma^{4\varphi} \tau_\varphi^{2N(K-1)} (\mathbb{E}[\eta_t^2])^2 \right] , \\
&= \sigma^{4\varphi} e^{2\xi} \tau_\varphi^{2N(K-1)} \left(\psi_\varphi^{N(K-1)} \mathbb{E}[\eta_t^2] - (\mathbb{E}[\eta_t^2])^2 \right) .
\end{aligned}$$

Pour prouver les résultats sur les espérances des fonctions de ϵ_t et de ε_t , nous avons besoin d'un résultat intermédiaire présenté dans le lemme suivant :

Lemme 2. Soit X une variable aléatoire normale centrée réduite, λ_1 un nombre réel et $\lambda_2 < \frac{1}{2}$. On a :

$$\begin{aligned}
\mathbb{E}[\exp(\lambda_1 X)] &= \exp \left(\frac{\lambda_1^2}{2} \right) , \\
\mathbb{E}[\exp(\lambda_2 X^2 + \lambda_1 X)] &= \frac{1}{\sqrt{1-2\lambda_2}} \exp \left(\frac{\lambda_1^2}{2(1-2\lambda_2)} \right) .
\end{aligned}$$

DÉMONSTRATION. La fonction génératrice des moments d'une variable aléatoire X distribuée selon une loi normale de moyenne μ et variance σ^2 est donnée par

$$\lambda \mapsto \mathbb{E}[\exp(\lambda X)] = \exp \left(\mu \lambda + \frac{\sigma^2 \lambda^2}{2} \right) .$$

Donc pour $\mu = 0$ et $\sigma = 1$ on a

$$\mathbb{E}[\exp(\lambda_1 X)] = \exp \left(\frac{\lambda_1^2}{2} \right) .$$

La fonction génératrice des moments d'une variable aléatoire Y distribuée selon une loi du khi deux non centré de centre μ est donnée par

$$\lambda \mapsto \mathbb{E}[\exp(\lambda Y)] = \frac{\exp \left(\frac{\mu^2 \lambda}{1-2\lambda} \right)}{\sqrt{1-2\lambda}} ,$$

si $\lambda < \frac{1}{2}$. Donc pour $\mu = \frac{\lambda_1}{2\lambda_2}$ on a

$$\mathbb{E}[\exp(\lambda_2 Y)] = \frac{\exp\left(\frac{\left(\frac{\lambda_1}{2\lambda_2}\right)^2 \lambda_2}{1-2\lambda_2}\right)}{\sqrt{1-2\lambda_2}} = \frac{\exp\left(\frac{\lambda_1^2}{4\lambda_2(1-2\lambda_2)}\right)}{\sqrt{1-2\lambda_2}}.$$

Ainsi, par une complétion de carré on a :

$$\begin{aligned} \mathbb{E}[\exp(\lambda_2 X^2 + \lambda_1 X)] &= \mathbb{E}\left[\exp\left(\lambda_2 \left(X + \frac{\lambda_1}{2\lambda_2}\right)^2 - \frac{\lambda_1^2}{4\lambda_2}\right)\right], \\ &= \mathbb{E}\left[\exp\left(\lambda_2 \left(X + \frac{\lambda_1}{2\lambda_2}\right)^2\right)\right] \exp\left(-\frac{\lambda_1^2}{4\lambda_2}\right), \\ &= \mathbb{E}[\exp(\lambda_2 Y)] \exp\left(-\frac{\lambda_1^2}{4\lambda_2}\right), \end{aligned}$$

où $Y = \left(X + \frac{\lambda_1}{2\lambda_2}\right)^2$. Puisque X est distribuée selon une loi normale centrée réduite, Y est distribuée selon une loi du khi deux non centré de centre $\frac{\lambda_1^2}{2\lambda_2}$. Ainsi,

$$\begin{aligned} \mathbb{E}[\exp(\lambda_2 X^2 + \lambda_1 X)] &= \frac{\exp\left(\frac{\lambda_1^2}{4\lambda_2(1-2\lambda_2)}\right)}{\sqrt{1-2\lambda_2}} \exp\left(-\frac{\lambda_1^2}{4\lambda_2}\right), \\ &= \frac{1}{\sqrt{1-2\lambda_2}} \exp\left(\frac{\lambda_1^2}{4\lambda_2(1-2\lambda_2)} - \frac{\lambda_1^2}{4\lambda_2}\right), \\ &= \frac{1}{\sqrt{1-2\lambda_2}} \exp\left(\frac{\lambda_1^2}{2(1-2\lambda_2)}\right). \end{aligned}$$

□

Puisque $\eta_t = \exp(\delta_1 \epsilon_t + \delta_2 \epsilon_t^2 + \gamma \epsilon_t - \delta_2)$ et que les processus $\{\epsilon_t\}$ et $\{\varepsilon_t\}$ sont normalement distribués, on a :

$$\begin{aligned} \mathbb{E}[\eta_t] &= \mathbb{E}[\exp(\delta_1 \epsilon_t + \delta_2 \epsilon_t^2 + \gamma \epsilon_t - \delta_2)], \\ &= \exp(-\delta_2) \mathbb{E}[\exp(\delta_1 \epsilon_t + \delta_2 \epsilon_t^2)] \mathbb{E}[\gamma \epsilon_t], \\ &= \exp(-\delta_2) \frac{1}{\sqrt{1-2\delta_2}} \exp\left(\frac{\delta_1^2}{2(1-2\delta_2)}\right) \exp\left(\frac{\gamma^2}{2}\right), \\ &= \frac{1}{\sqrt{1-2\delta_2}} \exp\left(\frac{1}{2}\left(\frac{\delta_1^2}{1-2\delta_2} + \gamma^2\right) - \delta_2\right). \end{aligned}$$

Aussi,

$$\begin{aligned} \mathbb{E}[\eta_t^2] &= \mathbb{E}[\exp(2\delta_1 \epsilon_t + 2\delta_2 \epsilon_t^2 + 2\gamma \epsilon_t - 2\delta_2)], \\ &= \frac{1}{\sqrt{1-4\delta_2}} \exp\left(\frac{2\delta_1^2}{1-2\delta_2} + 2\gamma^2 - 2\delta_2\right). \end{aligned}$$

□

A.7. Preuve des résultats de prévision

Soit (Y_t, V_t) un modèle à chaîne de Markov cachée dans lequel le processus $\{Y_t\}$ est régi par l'une des équations (2.3.1), (2.3.2), (2.3.3) ou (2.3.4). Puisque $\{V_t\}$ est une chaîne de Markov homogène à espace d'états de dimension fini (K^N), les moments conditionnels d'une fonction puissance de V_{t+h} sont donnés par :

$$\mathbb{E} [V_{t+h}^\varphi \mid Y_{1:t}] = \sum_{j=1}^{K^N} \mathbb{P}(V_{t+h} = \nu_j \mid Y_{1:t}) \nu_j^\varphi,$$

où φ est une constante réelle. (voir Hamilton, 1994, chapitre 22 ou Zucchini et MacDonald, 2009, chapitre 2). Désignons respectivement par $\xi_{t|t}$ et $\xi_{t+h|t}$ le vecteur de probabilités filtrées et le vecteur des probabilités conditionnelles de prévision, c'est-à-dire pour tout $h = 0, 1, \dots$,

$$\xi_{j,t+h|t} = \mathbb{P}(V_{t+h} = \nu_j \mid Y_{1:t}), \quad j = 1, 2, \dots, K^N,$$

sont les composantes du vecteur $\xi_{t+h|t}$. On a alors

$$\mathbb{E} [V_{t+h}^\varphi \mid Y_{1:t}] = \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j^\varphi.$$

De plus, on a

$$\begin{aligned} \mathbb{P}(V_{t+h} = \nu_j \mid Y_{1:t}) &= \sum_{i=1}^{K^N} \mathbb{P}(V_{t+h} = \nu_j, V_t = \nu_i \mid Y_{1:t}), \\ &= \sum_{i=1}^{K^N} \mathbb{P}(V_{t+h} = \nu_j \mid Y_{1:t}, V_t = \nu_i) \mathbb{P}(V_t = \nu_i \mid Y_{1:t}), \\ &= \sum_{i=1}^{K^N} \mathbb{P}(V_{t+h} = \nu_j \mid V_t = \nu_i) \mathbb{P}(V_t = \nu_i \mid Y_{1:t}). \end{aligned}$$

La représentation vectorielle du résultat,

$$\mathbb{P}(V_{t+h} = \nu_j \mid Y_{1:t}) = \sum_{i=1}^{K^N} \mathbb{P}(V_{t+h} = \nu_j \mid V_t = \nu_i) \mathbb{P}(V_t = \nu_i \mid Y_{1:t}),$$

est la relation

$$\xi_{t+h|t} = \mathbf{P}^h \xi_{t|t}, \quad h = 1, 2, \dots,$$

où \mathbf{P} désigne la matrice de transition de la chaîne de Markov $\{V_t\}$.

Nous supposons que le processus $\{Y_t\}$ est régi par l'équation

$$Y_t = e^{\xi} V_t^\varphi \eta_t. \tag{A.7.1}$$

Les équations (2.3.1), (2.3.2), (2.3.3) et (2.3.4) peuvent être vues comme des cas particulier de l'équation (A.7.1). Cette dernière sera donc utilisée pour déduire les résultats concernant les autres équations. On a

$$\begin{aligned}\mathbb{E}[Y_{t+h} | Y_{1:t}] &= \mathbb{E}[e^{\xi} V_{t+h}^{\varphi} \eta_{t+h} | Y_{1:t}], \\ &= e^{\xi} \mathbb{E}[V_{t+h}^{\varphi} | Y_{1:t}] \mathbb{E}[\eta_{t+h}],\end{aligned}$$

puisque les processus $\{\eta_t\}$ et $\{V_t\}$ sont indépendants et que η_{t+h} est indépendant de $Y_{1:t}$. Par suite,

$$\mathbb{E}[Y_{t+h} | Y_{1:t}] = e^{\xi} \mathbb{E}[\eta_{t+h}] \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j^{\varphi}.$$

Si le processus $\{Y_t\}$ est régi par l'équation (2.3.1) ou l'équation (2.3.3), c'est-à-dire $Y_t = r_t^2$, alors $\xi = 0$, $\varphi = 1$ et $\eta_t = \epsilon_t^2$. On a donc

$$\mathbb{E}[r_{t+h}^2 | Y_{1:t}] = \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j.$$

Si le processus $\{Y_t\}$ est régi par l'équation (2.3.2), c'est-à-dire $Y_t = RV_t$, alors $\xi = 0$, $\varphi = 1$ et $\mathbb{E}[\eta_t] = 1$.

Donc

$$\mathbb{E}[RV_{t+h} | Y_{1:t}] = \sum_{j=1}^{K^N} \xi_{j,t+h|t} \nu_j.$$

Si le processus $\{Y_t\}$ est régi par l'équation (2.3.4) alors $\eta_t = \exp(\delta_1 \epsilon_t + \delta_2 \epsilon_t^2 + \gamma \epsilon_t - \delta_2)$. D'après la proposition 2.3.1, si les processus $\{\epsilon_t\}$ et $\{\varepsilon_t\}$ sont normalement distribués, alors

$$\mathbb{E}[\eta_t] = \frac{1}{\sqrt{1 - 2\delta_2}} \exp\left(\frac{1}{2} \left(\frac{\delta_1^2}{1 - 2\delta_2} + \gamma^2\right) - \delta_2\right).$$

Annexe B

Annexe relative au chapitre 3

B.1. Estimation et prévision dans le modèle de Liu et Maheu (2018)

Le modèle introduit par Liu et Maheu (2018) est un modèle joint qui relie les rendements financiers aux variances réalisées à travers une variable latente qui est une chaîne de Markov $\{s_t\}$ à espace d'états $\{1, 2, \dots, K\}$. Il modélise le processus joint de rendements financiers et de variance réalisées $\{(r_t, RV_t)\}$ comme suit :

$$\begin{cases} r_t | s_t \sim \text{Normal}(\mu_{s_t}, \exp(\sigma_{s_t})), \\ RV_t | s_t \sim \text{Inv-Gamma}(\nu + 1, \nu\sigma_{s_t}^2), \end{cases}$$

où $\nu, \mu_1, \sigma_1, \mu_2, \sigma_2, \dots, \mu_K, \sigma_K$ sont des paramètres à estimer. Liu et Maheu (2018) introduisent des alternatives à ce modèle que nous ne présenterons pas dans cette thèse.

Dans cette section, nous présentons les résultats d'estimation et de prévision du modèle de Liu et Maheu (2018) en supposant que $\mu_1 = \mu_2 = \dots = \mu_K = 0$ que nous comparons au modèle MDSV. Le tableau B.1 présente les résultats d'estimation du modèle de Liu et Maheu (2018) et des trois spécifications du modèle MDSV.

Le tableau B.2 présente les log-vraisemblances prédictives du modèle de Liu et Maheu (2018) et des trois spécifications du modèle MDSV.

Le tableau B.3 présente les écarts de prévision quadratique et absolu, ainsi que la quasi-vraisemblance du modèle de Liu et Maheu (2018) et des trois spécifications du modèle MDSV aux horizons $h = 1, 5, 25, 100$ jours.

B.2. Comparaison de prévision de variances réalisées : forme univariée ou forme jointe

D'après le tableau B.4, la forme jointe produit de meilleures prévisions que la forme univariée (90 cas sur 144). Plus particulièrement, à court terme (une semaine ou moins), dans 55 cas sur 72, la forme jointe du modèle MDSV produit de meilleures prévisions que la forme univariée. Par exemple, en dehors de la spécification ($N = 3, K = 10$) ajusté sur le NASDAQ 100, la forme jointe produit des RMSFE, MAFE et QLIK à l'horizon d'un jour plus faibles que la forme univariée. De plus, le test de Diebold-Mariano indique que les MAFE à l'horizon d'un jour produits par la forme jointe des trois spécifications sont significativement

TABLE B.1 – Comparaison de performance : Modèle de Liu et Maheu (2018) et MDSV

| Modèles | Sans effet levier | | | | | Avec effet levier | | | | |
|-----------------------------------|-------------------|----------|---------------------|--------------------|---------------------|-------------------|----------|---------------------|--------------------|---------------------|
| | Modèle de Maheu | | MDSV | | | Modèle de Maheu | | MDSV | | |
| | $K = 2$ | $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | $K = 2$ | $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| Np | 5 | 17 | 10 | 10 | 10 | 7 | 19 | 12 | 12 | 12 |
| S&P 500 ($n = 5016$) | | | | | | | | | | |
| \mathcal{L} | -9550.7 | -7993.3 | -6979.0 | -6977.0 | -6996.0 | -8324.1 | -7402.2 | -6673.8 | -6687.9 | -6701.0 |
| AIC | -9555.7 | -8010.3 | -6989.0 | -6987.0 | -7006.0 | -8331.1 | -7421.2 | -6685.8 | -6699.9 | -6713.0 |
| BIC | -9572.0 | -8065.7 | -7021.6 | -7019.6 | -7038.6 | -8354.0 | -7483.2 | -6724.9 | -6739.0 | -6752.2 |
| NASDAQ 100 ($n = 5012$) | | | | | | | | | | |
| \mathcal{L} | -11512.4 | -10187.1 | -8896.5 | -8892.0 | -8914.5 | -10624.0 | -9802.2 | -8546.9 | -8568.3 | -8590.3 |
| AIC | -11517.4 | -10204.1 | -8906.5 | -8902.0 | -8924.5 | -10631.0 | -9821.2 | -8558.9 | -8580.3 | -8602.3 |
| BIC | -11533.7 | -10259.5 | -8939.1 | -8934.6 | -8957.1 | -10653.8 | -9883.2 | -8598.0 | -8619.4 | -8641.5 |
| FTSE 100 ($n = 5042$) | | | | | | | | | | |
| \mathcal{L} | -9990.6 | -8686.6 | -8044.7 | -8037.3 | -8050.7 | -8918.2 | -8302.4 | -7761.4 | -7804.8 | -7789.4 |
| AIC | -9995.6 | -8703.6 | -8054.7 | -8047.3 | -8060.7 | -8925.2 | -8321.4 | -7773.4 | -7816.8 | -7801.4 |
| BIC | -10011.9 | -8759.0 | -8087.3 | -8079.9 | -8093.3 | -8948.0 | -8383.4 | -7812.6 | -7855.9 | -7840.6 |
| NIKKEI 225 ($n = 4865$) | | | | | | | | | | |
| \mathcal{L} | -11394.8 | -10334.3 | -9059.8 | -9076.1 | -9122.9 | -10586.1 | -10000.5 | -8839.2 | -8863.2 | -8877.5 |
| AIC | -11399.8 | -10351.3 | -9069.8 | -9086.1 | -9132.9 | -10593.1 | -10019.5 | -8851.2 | -8875.2 | -8889.5 |
| BIC | -11416.1 | -10406.5 | -9102.3 | -9118.6 | -9165.3 | -10615.8 | -10081.1 | -8890.1 | -8914.1 | -8928.4 |

Np : Nombre de paramètres, \mathcal{L} : log-vraisemblance, AIC : Critère d'information d'Akaike, BIC : Critère d'information bayésien, n : Taille d'échantillon. Les valeurs les plus élevées sont en gras.

TABLE B.2 – Log-vraisemblance prédictive : Modèle de Liu et Maheu (2018) et MDSV

| Modèles | Sans effet levier | | | | | Avec effet levier | | | | |
|------------|-------------------|---------|---------------------|--------------------|---------------------|-------------------|---------|---------------------|--------------------|---------------------|
| | Modèle de Maheu | | MDSV | | | Modèle de Maheu | | MDSV | | |
| | $K = 2$ | $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) | $K = 2$ | $K = 4$ | (N, K) (3, 10) | (N, K) (6, 3) | (N, K) (10, 2) |
| S&P 500 | -778.6 | -731.9 | -711.5 | -713.2 | -718.3 | -732.1 | -716.5 | -708.0* | -711.2 | -712.7 |
| NASDAQ 100 | -973.3 | -948.0 | -924.5 | -924.2 | -923.2* | -946.5 | -948.2 | -924.1 | -923.4 | -924.2 |
| FTSE 100 | -808.9 | -787.8 | -790.5 | -792.0 | -792.7 | -788.9 | -783.8 | -782.4* | -789.5 | -783.8 |
| NIKKEI 225 | -990.3 | -997.8 | -975.5 | -978.0 | -978.7 | -980.0 | -1003.0 | -973.7* | -975.1 | -975.4 |

Les valeurs les plus élevées sont en gras. L'étoile (*) marque le meilleur modèle MDSV.

(1%) plus faibles que ceux produits par la forme univariée sur les données du NASDAQ 100 et du FTSE 100. À long terme (plus d'une semaine), la forme univariée produit de meilleures prédictions que la forme jointe environ la moitié du temps (37 cas sur 70).

Ces résultats concordent aussi avec la littérature financière. Ainsi, modéliser simultanément les rendements et les variances réalisées permet de mieux accéder à la volatilité et donc de faire de meilleures prévisions, surtout à court terme. Cependant, les prévisions à long terme des variances réalisées par un modèle joint peuvent être très affectées par l'incertitude des prévisions des rendements financiers.

TABLE B.3 – Performance de prévision de volatilité : Modèle de [Liu et Maheu \(2018\)](#) et MDSV

| | | RMSFE | | | | MAFE | | | | QLIK | | | |
|--------------------|------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|
| horizon(h) | | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 |
| S&P 500 | | | | | | | | | | | | | |
| Sans effet levier | Maheu($K = 2$) | 0.68 | 0.65 | 0.69 | 0.77 | 0.46 | 0.49 | 0.63 | 0.71 | -0.06 | 1.68 | 3.54 | 5.13 |
| | Maheu($K = 4$) | 0.54 | 0.53 | 0.57 | 0.65 | 0.28 | 0.32 | 0.50 | 0.59 | -0.27 | 1.49 | 3.42 | 5.05 |
| | MDSV(3, 10) | 0.48 | 0.45 | 0.44 | 0.45 | 0.22 | 0.23 | 0.32 | 0.39 | -0.37 | 1.38 | 3.26 | 4.88 |
| | MDSV(6, 3) | 0.48 | 0.43 | 0.40 | 0.37 | 0.21 | 0.21 | 0.27 | 0.31 | -0.37 | 1.37 | 3.22 | 4.91 |
| | MDSV(10, 2) | 0.49 | 0.43 | 0.39 | 0.35 | 0.21 | 0.21 | 0.26 | 0.28 | -0.37 | 1.37 | 3.26 | 4.89 |
| Avec effet levier | Maheu($K = 2$) | 0.52 | 0.50 | 0.52 | 0.56 | 0.30 | 0.33 | 0.46 | 0.50 | -0.25 | 1.48 | 3.39 | 5.00 |
| | Maheu($K = 4$) | 0.63 | 0.77 | 0.91 | 1.67 | 0.26 | 0.30 | 0.48 | 0.97 | -0.35 | 1.41 | 3.37 | 5.13 |
| | MDSV(3, 10) | 0.45 | 0.42 | 0.41 | 0.37 | 0.21 | 0.21 | 0.27 | 0.32 | -0.37 | 1.35 | 3.23 | 4.83 |
| | MDSV(6, 3) | 0.44 | 0.41 | 0.40 | 0.35 | 0.20 | 0.19 | 0.23 | 0.25 | -0.38 | 1.35 | 3.26 | 4.96 |
| | MDSV(10, 2) | 0.43 | 0.41 | 0.39 | 0.35 | 0.19 | 0.19 | 0.23 | 0.25 | -0.38 | 1.36 | 3.28 | 5.03 |
| NASDAQ 100 | | | | | | | | | | | | | |
| Sans effet levier | Maheu($K = 2$) | 1.09 | 1.06 | 1.08 | 1.22 | 0.73 | 0.76 | 0.92 | 1.13 | 0.35 | 2.04 | 3.85 | 5.47 |
| | Maheu($K = 4$) | 0.74 | 0.79 | 0.83 | 0.91 | 0.43 | 0.48 | 0.65 | 0.81 | 0.13 | 1.85 | 3.69 | 5.33 |
| | MDSV(3, 10) | 0.60 | 0.51 | 0.45 | 0.41 | 0.27 | 0.28 | 0.32 | 0.34 | 0.03 | 1.74 | 3.52 | 5.08 |
| | MDSV(6, 3) | 0.59 | 0.51 | 0.45 | 0.41 | 0.27 | 0.27 | 0.31 | 0.35 | 0.02 | 1.73 | 3.53 | 5.10 |
| | MDSV(10, 2) | 0.59 | 0.51 | 0.44 | 0.39 | 0.27 | 0.26 | 0.29 | 0.31 | 0.02 | 1.73 | 3.52 | 5.09 |
| Avec effet levier | Maheu($K = 2$) | 0.86 | 5.44 | 4.09 | 5.20 | 0.55 | 0.87 | 1.07 | 1.13 | 0.20 | 1.93 | 3.76 | 5.35 |
| | Maheu($K = 4$) | 0.69 | 0.70 | 0.84 | 0.74 | 0.34 | 0.36 | 0.46 | 0.59 | 0.04 | 1.75 | 3.58 | 5.22 |
| | MDSV(3, 10) | 0.55 | 0.49 | 0.44 | 0.41 | 0.25 | 0.25 | 0.30 | 0.33 | 0.02 | 1.73 | 3.53 | 5.10 |
| | MDSV(6, 3) | 0.55 | 0.49 | 0.44 | 0.39 | 0.25 | 0.25 | 0.30 | 0.31 | 0.02 | 1.73 | 3.52 | 5.08 |
| | MDSV(10, 2) | 0.55 | 0.49 | 0.44 | 0.39 | 0.24 | 0.25 | 0.28 | 0.29 | 0.02 | 1.73 | 3.52 | 5.11 |
| FTSE 100 | | | | | | | | | | | | | |
| Sans effet levier | Maheu($K = 2$) | 0.64 | 0.47 | 0.49 | 0.64 | 0.36 | 0.34 | 0.44 | 0.62 | 0.20 | 1.85 | 3.58 | 5.11 |
| | Maheu($K = 4$) | 0.56 | 0.37 | 0.38 | 0.56 | 0.26 | 0.23 | 0.33 | 0.53 | 0.13 | 1.79 | 3.51 | 5.06 |
| | MDSV(3, 10) | 0.56 | 0.35 | 0.28 | 0.32 | 0.24 | 0.19 | 0.21 | 0.27 | 0.13 | 1.78 | 3.44 | 4.92 |
| | MDSV(6, 3) | 0.54 | 0.32 | 0.23 | 0.21 | 0.23 | 0.18 | 0.18 | 0.18 | 0.13 | 1.77 | 3.43 | 4.86 |
| | MDSV(10, 2) | 0.54 | 0.31 | 0.22 | 0.18 | 0.23 | 0.17 | 0.16 | 0.14 | 0.13 | 1.77 | 3.43 | 4.85 |
| Avec effet levier | Maheu($K = 2$) | 0.57 | 0.39 | 0.37 | 0.52 | 0.28 | 0.24 | 0.32 | 0.50 | 0.13 | 1.77 | 3.50 | 5.04 |
| | Maheu($K = 4$) | 0.56 | 0.37 | 0.37 | 0.54 | 0.26 | 0.21 | 0.28 | 0.48 | 0.12 | 1.75 | 3.46 | 5.03 |
| | MDSV(3, 10) | 0.57 | 0.33 | 0.22 | 0.19 | 0.24 | 0.17 | 0.16 | 0.15 | 0.12 | 1.75 | 3.42 | 4.85 |
| | MDSV(6, 3) | 0.54 | 0.31 | 0.22 | 0.21 | 0.22 | 0.16 | 0.16 | 0.18 | 0.11 | 1.76 | 3.43 | 4.86 |
| | MDSV(10, 2) | 0.54 | 0.31 | 0.21 | 0.17 | 0.23 | 0.16 | 0.15 | 0.13 | 0.11 | 1.75 | 3.43 | 4.86 |
| NIKKEI 225 | | | | | | | | | | | | | |
| Sans effet levier | Maheu($K = 2$) | 0.72 | 0.73 | 0.91 | 1.17 | 0.58 | 0.63 | 0.86 | 1.15 | 0.23 | 1.91 | 3.73 | 5.31 |
| | Maheu($K = 4$) | 0.53 | 0.53 | 0.63 | 0.85 | 0.32 | 0.38 | 0.56 | 0.82 | -0.03 | 1.69 | 3.54 | 5.14 |
| | MDSV(3, 10) | 0.42 | 0.33 | 0.29 | 0.30 | 0.20 | 0.18 | 0.22 | 0.26 | -0.11 | 1.54 | 3.28 | 4.78 |
| | MDSV(6, 3) | 0.41 | 0.32 | 0.26 | 0.23 | 0.19 | 0.17 | 0.18 | 0.20 | -0.11 | 1.54 | 3.27 | 4.75 |
| | MDSV(10, 2) | 0.41 | 0.31 | 0.25 | 0.21 | 0.19 | 0.16 | 0.18 | 0.19 | -0.12 | 1.54 | 3.27 | 4.74 |
| Avec effet levier | Maheu($K = 2$) | 0.56 | 0.54 | 0.57 | 0.71 | 0.38 | 0.40 | 0.50 | 0.68 | 0.03 | 1.70 | 3.49 | 5.06 |
| | Maheu($K = 4$) | 0.57 | 0.57 | 0.78 | 0.82 | 0.32 | 0.35 | 0.50 | 0.67 | -0.05 | 1.64 | 3.46 | 5.02 |
| | MDSV(3, 10) | 0.41 | 0.31 | 0.26 | 0.24 | 0.20 | 0.18 | 0.20 | 0.22 | -0.13 | 1.53 | 3.27 | 4.75 |
| | MDSV(6, 3) | 0.39 | 0.30 | 0.25 | 0.21 | 0.19 | 0.16 | 0.18 | 0.19 | -0.13 | 1.53 | 3.26 | 4.74 |
| | MDSV(10, 2) | 0.39 | 0.30 | 0.24 | 0.20 | 0.19 | 0.17 | 0.18 | 0.18 | -0.13 | 1.52 | 3.26 | 4.73 |

B.3. Estimation du modèle MDSV par maximum de vraisemblance sur plusieurs séries de données

Dans cette section, nous présentons les paramètres du modèle MDSV joint estimés par maximum de vraisemblance sur les indices financiers S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225 à la travers le tableau [B.5](#).

B.4. Fonctions d'autocorrélation du modèle MDSV sur plusieurs séries de données

Dans cette section, nous présentons les fonctions d'autocorrélation du modèle MDSV sur les quatre indices financiers (S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225) à travers la figure [B.1](#).

TABLE B.4 – Prédiction de variances réalisées : forme univariée ou forme jointe (avec effet levier)

| horizon(h) | | RMSFE | | | | MAFE | | | | QLIK | | | | |
|--------------------|---|-----------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|--------------------------|-------------|--------------------------|
| | | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 | 1 | 5 | 25 | 100 | |
| S&P 500 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.46 | 0.41 | 0.39 | 0.33 | 0.21 | 0.21 | 0.27 | 0.28 | -0.36 | 1.36 | 3.23 | 4.83 |
| | | jointe | 0.45 | 0.42 | 0.41 | 0.37 | 0.21 | 0.21 | 0.27 | 0.32 | -0.37 | 1.35 [†] | 3.23 | 4.83 |
| MDSV (6, 3) | { | univariée | 0.44 | 0.41 | 0.39 | 0.34 | 0.20 | 0.20 | 0.26 | 0.27 | -0.38 | 1.36 | 3.26 | 4.94 |
| | | jointe | 0.44 | 0.41 | 0.40 | 0.35 | 0.20 * | 0.19 [†] | 0.23 [‡] | 0.25 | -0.38 | 1.35 | 3.26 | 4.96 |
| MDSV (10, 2) | { | univariée | 0.44 | 0.41 | 0.39 | 0.34 | 0.20 | 0.19 | 0.24 | 0.25 | -0.38 | 1.36 | 3.26 | 4.98 |
| | | jointe | 0.43 | 0.41 | 0.39 | 0.35 | 0.19 [†] | 0.19 | 0.23 [‡] | 0.25 | -0.38 [†] | 1.36 | 3.28 | 5.03 |
| NASDAQ 100 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.55 | 0.50 | 0.46 | 0.48 | 0.27 | 0.28 | 0.35 | 0.42 | 0.02 | 1.73 | 3.54 | 5.13 |
| | | jointe | 0.55 | 0.49 | 0.44 * | 0.41 [‡] | 0.25 [‡] | 0.25 [‡] | 0.30 [‡] | 0.33 [‡] | 0.02 | 1.73 | 3.53 | 5.10 |
| MDSV (6, 3) | { | univariée | 0.55 | 0.49 | 0.44 | 0.39 | 0.26 | 0.25 | 0.28 | 0.29 | 0.02 | 1.73 | 3.54 | 5.17 |
| | | jointe | 0.55 | 0.49 | 0.44 | 0.39 | 0.25 [‡] | 0.25 | 0.30 | 0.31 | 0.02 | 1.73 | 3.52 | 5.08 |
| MDSV (10, 2) | { | univariée | 0.55 | 0.49 | 0.44 | 0.38 | 0.26 | 0.25 | 0.27 | 0.27 | 0.01 | 1.73 | 3.53 | 5.13 |
| | | jointe | 0.55 | 0.49 | 0.44 | 0.39 | 0.24 [‡] | 0.25 | 0.28 | 0.29 | 0.02 | 1.73 | 3.52 | 5.11 |
| FTSE 100 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.66 | 0.34 | 0.25 | 0.25 | 0.26 | 0.18 | 0.20 | 0.22 | 0.14 | 1.76 | 3.44 | 4.89 |
| | | jointe | 0.57 | 0.33 [†] | 0.22 [†] | 0.19 [‡] | 0.24 [‡] | 0.17 [†] | 0.16 [‡] | 0.15 [‡] | 0.12 [‡] | 1.75 [†] | 3.42 | 4.85 [†] |
| MDSV (6, 3) | { | univariée | 0.64 | 0.32 | 0.22 | 0.22 | 0.26 | 0.18 | 0.18 | 0.18 | 0.13 | 1.74 | 3.42 | 4.88 |
| | | jointe | 0.54 | 0.31 | 0.22 | 0.21 | 0.22 [‡] | 0.16 [†] | 0.16 [†] | 0.18 | 0.11 | 1.76 | 3.43 | 4.86 |
| MDSV (10, 2) | { | univariée | 0.61 | 0.31 | 0.21 | 0.17 | 0.25 | 0.17 | 0.17 | 0.14 | 0.14 | 1.76 | 3.42 | 4.84 |
| | | jointe | 0.54 * | 0.31 | 0.21 | 0.17 | 0.23 [‡] | 0.16 [†] | 0.15 [†] | 0.13 | 0.11 [‡] | 1.75 | 3.43 | 4.86 |
| NIKKEI 225 | | | | | | | | | | | | | | |
| MDSV (3, 10) | { | univariée | 0.48 | 0.32 | 0.25 | 0.24 | 0.21 | 0.17 | 0.19 | 0.22 | -0.09 | 1.53 | 3.25 | 4.76 |
| | | jointe | 0.41 | 0.31 | 0.26 | 0.24 | 0.20 * | 0.18 | 0.20 | 0.22 | -0.13 [‡] | 1.53 | 3.27 | 4.75 |
| MDSV (6, 3) | { | univariée | 0.42 | 0.30 | 0.25 | 0.20 | 0.20 | 0.16 | 0.17 | 0.18 | -0.11 | 1.53 | 3.28 | 4.74 |
| | | jointe | 0.39 [†] | 0.30 | 0.25 | 0.21 | 0.19 [†] | 0.16 | 0.18 | 0.19 | -0.13 [‡] | 1.53 | 3.26 | 4.74 |
| MDSV (10, 2) | { | univariée | 0.42 | 0.30 | 0.24 | 0.19 | 0.20 | 0.15 | 0.17 | 0.16 | -0.12 | 1.53 | 3.26 | 4.71 |
| | | jointe | 0.39 * | 0.30 | 0.24 | 0.20 | 0.19 | 0.17 | 0.18 | 0.18 | -0.13 [‡] | 1.52 | 3.26 | 4.73 |

Les symboles *, [†] et [‡] signifient respectivement que le critère est significativement inférieur à celui de la forme univariée aux seuils de 10%, 5% et 1% en utilisant le test de Diebold-Mariano.

TABLE B.5 – Paramètres estimés par maximum de vraisemblance

| indices | Spécification | ω | a | b | σ^2 | ν_0 | x_i | φ | δ_1 | δ_2 | γ | l_1 | θ_l | |
|-------------------------|---------------|-----------------|--------|--------|------------|---------|--------|-----------|------------|------------|----------|--------|------------|--------|
| Sans effet levier | S&P 500 | $N = 3, K = 10$ | 0.3693 | 0.9947 | 6.5155 | 1.0150 | 0.8199 | -0.3615 | 0.9511 | -0.0712 | 0.0858 | 0.1675 | - | - |
| | | $N = 6, K = 3$ | 0.3557 | 0.9982 | 2.7533 | 0.8404 | 0.7223 | -0.3622 | 0.9525 | -0.0706 | 0.0862 | 0.1657 | - | - |
| | | $N = 10, K = 2$ | 0.3323 | 0.9991 | 1.9170 | 0.6631 | 0.6717 | -0.3633 | 0.9522 | -0.0701 | 0.0866 | 0.1660 | - | - |
| | NASDAQ 100 | $N = 3, K = 10$ | 0.3684 | 0.9933 | 6.5180 | 1.6622 | 0.8313 | -0.6681 | 0.8852 | -0.1169 | 0.0101 | 0.1229 | - | - |
| | | $N = 6, K = 3$ | 0.4821 | 0.9982 | 2.7756 | 1.8791 | 0.7243 | -0.6667 | 0.8847 | -0.1175 | 0.0086 | 0.1201 | - | - |
| | | $N = 10, K = 2$ | 0.4280 | 0.9991 | 1.8905 | 1.4375 | 0.6857 | -0.6744 | 0.8890 | -0.1176 | 0.0144 | 0.1283 | - | - |
| | FTSE 100 | $N = 3, K = 10$ | 0.3521 | 0.9946 | 5.1266 | 1.1738 | 0.8338 | -0.2020 | 0.9396 | -0.0353 | 0.1051 | 0.1864 | - | - |
| | | $N = 6, K = 3$ | 0.3301 | 0.9976 | 2.3335 | 0.8433 | 0.7646 | -0.2055 | 0.9398 | -0.0357 | 0.1089 | 0.1911 | - | - |
| | | $N = 10, K = 2$ | 0.3089 | 0.9997 | 2.0430 | 0.4441 | 0.7050 | -0.2077 | 0.9382 | -0.0359 | 0.1116 | 0.1946 | - | - |
| | NIKKEI 225 | $N = 3, K = 10$ | 0.2362 | 0.9930 | 7.1463 | 1.5534 | 0.8467 | -0.8581 | 1.0488 | -0.071 | 0.0524 | 0.1179 | - | - |
| | | $N = 6, K = 3$ | 0.2977 | 0.9989 | 3.2251 | 0.9054 | 0.7662 | -0.8605 | 1.0489 | -0.0726 | 0.0547 | 0.1222 | - | - |
| | | $N = 10, K = 2$ | 0.2964 | 0.9998 | 2.2904 | 0.5087 | 0.7191 | -0.8724 | 1.0551 | -0.0735 | 0.0624 | 0.1323 | - | - |
| Avec effet levier | S&P 500 | $N = 3, K = 10$ | 0.3151 | 0.9989 | 24.1159 | 0.3156 | 0.8593 | -0.3848 | 0.9671 | -0.1171 | 0.0826 | 0.1381 | 0.7679 | 0.8875 |
| | | $N = 6, K = 3$ | 0.2241 | 0.9996 | 5.3535 | 0.2033 | 0.7686 | -0.3811 | 0.9615 | -0.1165 | 0.0793 | 0.1321 | 0.7357 | 0.8701 |
| | | $N = 10, K = 2$ | 0.2140 | 0.9998 | 2.6887 | 0.1809 | 0.7506 | -0.3896 | 0.9634 | -0.1156 | 0.0870 | 0.1476 | 0.7324 | 0.8658 |
| | NASDAQ 100 | $N = 3, K = 10$ | 0.3474 | 0.9973 | 12.7479 | 0.7170 | 0.8630 | -0.7109 | 0.9045 | -0.1711 | 0.0177 | 0.1072 | 0.4728 | 0.8284 |
| | | $N = 6, K = 3$ | 0.3210 | 0.9995 | 3.8608 | 0.7603 | 0.772 | -0.7132 | 0.9030 | -0.1692 | 0.0210 | 0.1153 | 0.4225 | 0.7858 |
| | | $N = 10, K = 2$ | 0.2476 | 0.9994 | 2.0578 | 0.7603 | 0.7373 | -0.7171 | 0.9021 | -0.1684 | 0.0257 | 0.1272 | 0.4201 | 0.7857 |
| | FTSE 100 | $N = 3, K = 10$ | 0.0960 | 0.9978 | 26.9892 | 0.2459 | 0.8273 | -0.1939 | 0.9674 | -0.0668 | 0.0868 | 0.1447 | 0.5185 | 0.9058 |
| | | $N = 6, K = 3$ | 0.4470 | 0.9861 | 1.0016 | 0.4236 | 0.8418 | -0.2457 | 0.9667 | -0.0670 | 0.1284 | 0.2129 | 0.4567 | 0.8722 |
| | | $N = 10, K = 2$ | 0.1286 | 0.9993 | 2.6743 | 0.2012 | 0.7769 | -0.2057 | 0.9716 | -0.0661 | 0.0980 | 0.1613 | 0.4954 | 0.8991 |
| | NIKKEI 225 | $N = 3, K = 10$ | 0.1186 | 0.9965 | 12.6735 | 0.6451 | 0.8356 | -0.9071 | 1.0886 | -0.1115 | 0.0629 | 0.1112 | 0.2577 | 0.8635 |
| | | $N = 6, K = 3$ | 0.1972 | 0.9961 | 11.6541 | 0.5653 | 0.8279 | -0.9132 | 1.0929 | -0.1125 | 0.0655 | 0.1149 | 0.2667 | 0.8808 |
| | | $N = 10, K = 2$ | 0.1510 | 0.9991 | 2.2346 | 0.4807 | 0.7741 | -0.9128 | 1.0879 | -0.1137 | 0.0680 | 0.1229 | 0.2763 | 0.8846 |

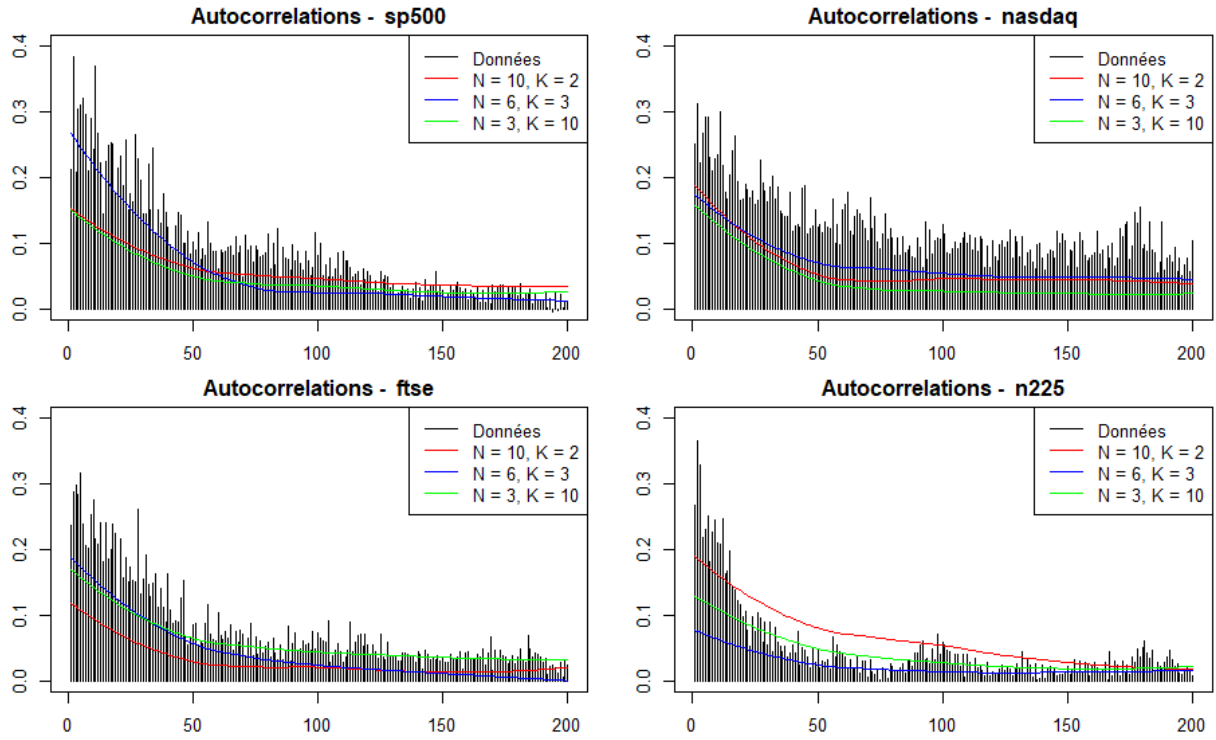


FIGURE B.1 – Fonction d'autocorrélation empirique en noir contre les fonctions d'autocorrélation théoriques des spécifications $(N = 10, K = 2)$ en rouge, $(N = 6, K = 3)$ en bleu et $(N = 3, K = 10)$ en vert du modèle MDSV sur les données des indices S&P 500, NASDAQ 100, FTSE 100 et NIKKEI 225.

Annexe C

Annexe relative au chapitre 4

C.1. Preuves de la proposition 4.2.1

Proposition.

(i) Conditionnellement aux probabilités p_1, \dots, p_N , la loi de \mathcal{S} est donnée par :

$$\mathbb{P}(\mathcal{S} \mid p_1, \dots, p_N) = \prod_{i=1}^N p_i^{n_{01}^{(i)} + n_{10}^{(i)}} (1 - p_i)^{n_{00}^{(i)} + n_{11}^{(i)}},$$

où $\mathbb{P}(\cdot)$ représente la loi.

(ii) Si on suppose que les probabilités p_i suivent une distribution bêta de paramètres $\frac{\alpha}{N}$ et 1 a priori, avec $\alpha > 0$, alors la loi marginale de \mathcal{S} est donnée par :

$$\mathbb{P}(\mathcal{S} \mid \alpha) = \prod_{i=1}^N \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}, \quad (\text{C.1.1})$$

où $\Gamma(\cdot)$ est la fonction Gamma et $n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} = T$ pour tout $i = 1, \dots, N$.

DÉMONSTRATION.

(i) Les chaînes de Markov $\{s_t^{(i)}\}$ sont indépendantes et vérifient $s_0^{(i)} = 0$. Donc pour chaque $\{s_t^{(i)}\}$, la probabilité d'avoir la colonne (i) comme réalisation est $p_i^{n_{01}^{(i)} + n_{10}^{(i)}} (1 - p_i)^{n_{00}^{(i)} + n_{11}^{(i)}}$. Le résultat s'obtient alors directement.

(ii) On a $p_i \sim \text{Beta}\left(\frac{\alpha}{N}, 1\right)$. Donc

$$\begin{aligned} \mathbb{P}(\mathcal{S} \mid \alpha) &= \int \cdots \int \mathbb{P}(\mathcal{S} \mid p_1, \dots, p_N) \prod_{i=1}^N \mathbb{P}(p_i \mid \alpha) dp_1 \dots dp_N, \\ &= \int \cdots \int \prod_{i=1}^N p_i^{n_{01}^{(i)} + n_{10}^{(i)}} (1 - p_i)^{n_{00}^{(i)} + n_{11}^{(i)}} \prod_{i=1}^N \mathbb{P}(p_i \mid \alpha) dp_1 \dots dp_N, \\ &= \prod_{i=1}^N \left[\int p_i^{n_{01}^{(i)} + n_{10}^{(i)}} (1 - p_i)^{n_{00}^{(i)} + n_{11}^{(i)}} \mathbb{P}(p_i \mid \alpha) dp_i \right]; \end{aligned}$$

en utilisant la densité de la loi beta, on obtient

$$\begin{aligned}
\mathbb{P}(\mathcal{S} \mid \alpha) &= \prod_{i=1}^N \left[\int p_i^{n_{01}^{(i)} + n_{10}^{(i)}} (1-p_i)^{n_{00}^{(i)} + n_{11}^{(i)}} \frac{\Gamma(\frac{\alpha}{N} + 1)}{\Gamma(\frac{\alpha}{N})\Gamma(1)} p_i^{\frac{\alpha}{N} - 1} dp_i \right], \\
&= \prod_{i=1}^N \frac{\Gamma(\frac{\alpha}{N} + 1)}{\Gamma(\frac{\alpha}{N})\Gamma(1)} \left[\int p_i^{n_{01}^{(i)} + n_{10}^{(i)} + \frac{\alpha}{N} - 1} (1-p_i)^{n_{00}^{(i)} + n_{11}^{(i)}} dp_i \right], \\
&= \prod_{i=1}^N \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}.
\end{aligned}$$

□

C.2. Preuves du théorème 4.2.1

Théorème (Loi de $[\mathcal{S}]$). Soit \mathcal{S} , une matrice binaire de dimension $T \times N$ dans laquelle chaque colonne d'indice i correspond aux états d'une chaîne de Markov $\{s_t^{(i)}\}$ qui prend ses valeurs dans $\{0, 1\}$ aux dates $t = 1, \dots, T$. On suppose que \mathcal{S} est une matrice lof; sinon, elle peut être transformée en matrice lof par une permutation des colonnes selon l'encodage binaire. Désignons aussi par N_h le nombre de colonnes de \mathcal{S} ayant le même encodage binaire h et par $[\mathcal{S}]$ la classe d'équivalence lof associée à \mathcal{S} . On a les résultats suivants :

(i) Le nombre d'éléments de la classe d'équivalence $[\mathcal{S}]$ est $\frac{N!}{\prod_{h=0}^{2^T-1} N_h!}$.

(ii) La loi de la classe d'équivalence $[\mathcal{S}]$ est

$$\mathbb{P}([\mathcal{S}] \mid \alpha) = \frac{N!}{\prod_{h=0}^{2^T-1} N_h!} \prod_{i=1}^N \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}. \quad (\text{C.2.1})$$

(iii) Lorsque $N \rightarrow +\infty$, la loi de la classe d'équivalence $[\mathcal{S}]$ est

$$\lim_{N \rightarrow +\infty} \mathbb{P}([\mathcal{S}] \mid \alpha) = \frac{\alpha^{N_+}}{\prod_{h=0}^{2^T-1} N_h!} \exp\{-\alpha H_T\} \prod_{i=1}^{N_+} \frac{\left(n_{01}^{(i)} + n_{10}^{(i)} - 1\right)! \left(n_{00}^{(i)} + n_{11}^{(i)}\right)!}{\left(n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)}\right)!}, \quad (\text{C.2.2})$$

où $H_T = \sum_{j=1}^T \frac{1}{j}$ et N_+ est le nombre de chaînes qui se sont activées au moins une fois entre $t = 1$ et $t = T$.

(iv) La dimension effective du modèle (N_+) est distribuée selon une loi de Poisson de paramètre αH_T .

DÉMONSTRATION.

(i) Le nombre d'éléments de la classe d'équivalence $[\mathcal{S}]$ correspond au nombre de permutations possibles des colonnes de \mathcal{S} . Si N_h représente le nombre de répétitions possibles de chacune des colonnes alors, le nombre de permutations est $\frac{N!}{\prod_{h=0}^{2^T-1} N_h!}$.

(ii) Toutes les matrices de la même classe ont la même probabilité. Ainsi, la probabilité de la classe d'équivalence est le nombre d'éléments de la classe, multiplié par la probabilité d'une matrice de cette classe. Par conséquent,

$$p([\mathcal{S}] | \alpha) = \frac{N!}{\prod_{h=0}^{2^T-1} N_h!} \prod_{i=1}^N \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}.$$

(iii) Avant de laisser tendre N vers l'infini, on a besoin de définir deux types de chaînes de Markov : celles qui ne s'activent jamais ($n_{00}^{(i)} = T, n_{10}^{(i)} = n_{01}^{(i)} = n_{11}^{(i)} = 0$) et les autres ($n_{00}^{(i)} < T$). Supposons qu'il y ait N_0 chaînes de Markov qui ne s'activent jamais et N_+ chaînes de Markov qui s'activent au moins une fois. On a alors $N = N_0 + N_+$. Comme \mathcal{S} est *lof*, les N_0 chaînes qui ne s'activent jamais sont positionnées après les N_+ chaînes qui s'activent au moins une fois. Ainsi,

$$\begin{aligned} p([\mathcal{S}] | \alpha) &= \frac{N!}{\prod_{h=0}^{2^T-1} N_h!} \prod_{i=1}^{N_+} \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)} \prod_{i=N_++1}^N \frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N}\right) \Gamma(T+1)}{\Gamma\left(\frac{\alpha}{N} + T + 1\right)}, \\ &= \frac{N!}{\prod_{h=0}^{2^T-1} N_h!} \left(\frac{\frac{\alpha}{N} \Gamma\left(\frac{\alpha}{N}\right) \Gamma(T+1)}{\Gamma\left(\frac{\alpha}{N} + T + 1\right)} \right)^{N_0} \left(\frac{\alpha}{N} \right)^{N_+} \prod_{i=1}^{N_+} \frac{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}, \\ &= \frac{N! \left(\frac{\alpha}{N}\right)^{N_+}}{\prod_{h=0}^{2^T-1} N_h!} \left(\frac{\Gamma\left(\frac{\alpha}{N} + 1\right) \Gamma(T+1)}{\Gamma\left(\frac{\alpha}{N} + T + 1\right)} \right)^N \prod_{i=1}^{N_+} \frac{\Gamma\left(\frac{\alpha}{N} + T + 1\right) \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)}\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma\left(\frac{\alpha}{N} + 1\right) \Gamma(T+1) \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}; \end{aligned}$$

En réexprimant certains ratio de fonctions Gamma à l'aide d'un produit de termes, nous obtenons

$$\begin{aligned} p([\mathcal{S}] | \alpha) &= \frac{N!}{\prod_{h=0}^{2^T-1} N_h!} \left(\frac{\alpha}{N} \right)^{N_+} \left(\frac{\Gamma(T+1)}{\prod_{t=1}^T \left(\frac{\alpha}{N} + t\right)} \right)^N \prod_{i=1}^{N_+} \frac{\Gamma\left(\frac{\alpha}{N} + T + 1\right) \prod_{k=1}^{n_{01}^{(i)} + n_{10}^{(i)} - 1} \left(\frac{\alpha}{N} + k\right) \Gamma\left(n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}{\Gamma(T+1) \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}, \\ &= \frac{\alpha^{N_+}}{\prod_{h=1}^{2^T-1} N_h!} \frac{N!}{N^{N_+} N_0!} \left(\frac{T!}{\prod_{t=1}^T \left(\frac{\alpha}{N} + t\right)} \right)^N \prod_{i=1}^{N_+} \frac{\Gamma\left(\frac{\alpha}{N} + T + 1\right) \prod_{k=1}^{n_{01}^{(i)} + n_{10}^{(i)} - 1} \left(\frac{\alpha}{N} + k\right) \left(n_{00}^{(i)} + n_{11}^{(i)}\right)!}{T! \Gamma\left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1\right)}. \end{aligned}$$

Remarquons ensuite que pour tout nombre entier M , on a

$$\lim_{N \rightarrow \infty} \prod_{k=1}^M \left(\frac{\alpha}{N} + k\right) = \lim_{N \rightarrow \infty} \left(M! + O\left(\frac{c}{N}\right)\right) = M!,$$

où c est une certaine constante. On a aussi

$$\begin{aligned} \lim_{N \rightarrow \infty} \left(\frac{T!}{\prod_{t=1}^T (t + \frac{\alpha}{N})} \right)^N &= \lim_{N \rightarrow \infty} \left(\prod_{t=1}^T \frac{t}{(t + \frac{\alpha}{N})} \right)^N = \lim_{N \rightarrow \infty} \prod_{t=1}^T \left(\frac{1}{1 + \frac{\alpha(\frac{1}{t})}{N}} \right)^N, \\ &= \prod_{t=1}^T \exp \left(-\alpha \left(\frac{1}{t} \right) \right) = \exp \left(-\alpha \sum_{t=1}^T \frac{1}{t} \right). \end{aligned}$$

Donc

$$\lim_{N \rightarrow \infty} \left(\frac{T!}{\prod_{t=1}^T (t + \frac{\alpha}{N})} \right)^N = \exp \left(-\alpha \sum_{t=1}^T \frac{1}{t} \right). \quad (\text{C.2.3})$$

Finalement,

$$\lim_{N \rightarrow \infty} \frac{N!}{N^{N+} N_0!} = \lim_{N \rightarrow \infty} \frac{\prod_{i=1}^{N_+} (N - i + 1)}{N^{N_+}} = \lim_{N \rightarrow \infty} \frac{N^{N_+} + O(N^{N_+})}{N^{N_+}} = 1.$$

Ainsi,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{\Gamma \left(\frac{\alpha}{N} + T + 1 \right) \prod_{k=1}^{n_{01}^{(i)} + n_{10}^{(i)} - 1} \left(\frac{\alpha}{N} + k \right)}{T! \Gamma \left(\frac{\alpha}{N} + n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} + 1 \right)} \\ = \lim_{N \rightarrow \infty} \frac{\Gamma \left(\frac{\alpha}{N} \right) \prod_{k=1}^T \left(\frac{\alpha}{N} + k \right) \prod_{k=1}^{n_{01}^{(i)} + n_{10}^{(i)} - 1} \left(\frac{\alpha}{N} + k \right)}{T! \Gamma \left(\frac{\alpha}{N} \right) \prod_{k=1}^{n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)}} \left(\frac{\alpha}{N} + k \right)}, \\ = \frac{T! \left(n_{01}^{(i)} + n_{10}^{(i)} - 1 \right)!}{T! \left(n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} \right)!}, \\ = \frac{\left(n_{01}^{(i)} + n_{10}^{(i)} - 1 \right)!}{\left(n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} \right)!}. \end{aligned}$$

Par la suite,

$$\lim_{N \rightarrow \infty} p([\mathcal{S}] | \alpha) = \frac{\alpha^{N_+}}{\prod_{h=1}^{2^T-1} N_h!} \exp \left(-\alpha \sum_{t=1}^T \frac{1}{t} \right) \prod_{i=1}^{N_+} \frac{\left(n_{01}^{(i)} + n_{10}^{(i)} - 1 \right)! \left(n_{00}^{(i)} + n_{11}^{(i)} \right)!}{\left(n_{01}^{(i)} + n_{10}^{(i)} + n_{00}^{(i)} + n_{11}^{(i)} \right)!}.$$

(iv) En se basant sur l'interprétation du processus du buffet indien présenté à la section 4.2.2, à chaque nouvelle ligne t , le client essaie un nombre de $\text{Poisson}(\alpha/t)$ de nouveaux plats. Donc le nombre effectif de colonnes est une somme de lois de Poisson indépendantes de moyenne $\alpha, \alpha/2, \dots, \alpha/T$. D'où N_+ est distribué selon une loi de Poisson de moyenne $\alpha \sum_{t=1}^T \frac{1}{t}$: $N_+ \sim \text{Poisson} \left(\alpha \sum_{t=1}^T \frac{1}{t} \right)$.

□

C.3. Preuves de la proposition 4.3.1

Proposition (Loi *a posteriori* de $p_{(N^*)}$). *La distribution a posteriori de la probabilité $p_{(N^*)}$ de la composante N^* pour laquelle les indices de toutes les composantes actives i vérifient $i < N^*$, est donnée par*

$$\mathbb{P}(p_{(N^*)} \mid p_{(N^*-1)}, \mathcal{S}_{:, \geq N^*} = 0) \propto \exp\left(\alpha \sum_{t=1}^T \frac{1}{t} (1 - p_{(N^*)})^t\right) p_{(N^*)}^{\alpha-1} (1 - p_{(N^*)})^T \mathbf{1}_{\{0 \leq p_{(N^*)} \leq p_{(N^*-1)}\}}, \quad (\text{C.3.1})$$

où $\mathcal{S}_{:, \geq N^*}$ représente la matrice composée de toutes les lignes de \mathcal{S} et de toutes les colonnes d'indice supérieur ou égal à N^* ; c'est une matrice de colonnes inactives.

DÉMONSTRATION. Tout d'abord, pour tout $i > 1$, posons l_i comme étant le nombre entier tel que $p_{l_i} = p_{(i)}$. En d'autres termes l_i est l'indice de la i^{eme} plus grande probabilité. Posons aussi $\mathbf{L}_i = \{1, \dots, N\} - \{l_1, \dots, l_i\}$. Il s'agit des indices des $N - i$ plus petites probabilités. Rappelons que les probabilités p_i , $i = 1, \dots, N$ sont indépendantes. Ainsi, conditionnellement à $p_{(1:i)}$, les probabilités p_l telles que $l \in \mathbf{L}_i$ sont indépendantes et identiquement distribuées. En utilisant les résultats sur les statistiques d'ordre, pour tout $l \in \mathbf{L}_i$, on a :

$$\mathbb{P}(p_l \mid p_{(1:i)}) = \mathbb{P}(p_l \mid p_{(i)}) = \frac{\alpha}{N} p_{(i)}^{-\frac{\alpha}{N}} p_l^{\frac{\alpha}{N}-1} \mathbf{1}_{\{0 < p_l < p_{(i)}\}}. \quad (\text{C.3.2})$$

De plus, pour $l \in \{1, \dots, N\}$, si $\mathcal{S}_{:,l}$ est la colonne d'indice l de \mathcal{S} on a :

$$\mathbb{P}(\mathcal{S}_{:,l} = 0 \mid p_l) = (1 - p_l)^T. \quad (\text{C.3.3})$$

Ainsi, si $\mathcal{S}_{:, \mathbf{L}_i}$ représente la matrice constituée des colonnes de \mathcal{S} d'indices l tels que $l \in \mathbf{L}_i$ et $p_{\mathbf{L}_i}$ alors :

$$\mathbb{P}(\mathcal{S}_{:, \mathbf{L}_i} = 0 \mid p_{\mathbf{L}_i}) \mathbb{P}(p_{\mathbf{L}_i} \mid p_{(i)}) = \prod_{l \in \mathbf{L}_i} (1 - p_l)^T \frac{\alpha}{N} p_{(i)}^{-\frac{\alpha}{N}} p_l^{\frac{\alpha}{N}-1} \mathbf{1}_{\{0 < p_l < p_{(i)}\}}.$$

Puisque $p_{(i:\infty)} = (p_{(k)})_{k=i, i+1, \dots}$ est la limite de $p_{\mathbf{L}_i} = \{p_l, l \in \mathbf{L}_i\}$ ordonné de façon décroissante lorsque $N \rightarrow \infty$, la probabilité que toutes les colonnes d'indice supérieur à (i) soient nulles conditionnellement à $p_{(i)}$ est

$$\begin{aligned} \mathbb{P}(\mathcal{S}_{:, > (i)} = 0 \mid p_{(i)}) &= \lim_{N \rightarrow \infty} \int \mathbb{P}(\mathcal{S}_{:, \mathbf{L}_i} = 0 \mid p_{\mathbf{L}_i}) \mathbb{P}(p_{\mathbf{L}_i} \mid p_{(i)}) dp_{\mathbf{L}_i}, \\ &= \lim_{N \rightarrow \infty} \left(\int_0^{p_{(i)}} (1 - p)^T \frac{\alpha}{N} p_{(i)}^{-\frac{\alpha}{N}} p^{\frac{\alpha}{N}-1} dp \right)^{N-i}. \end{aligned}$$

En effet, le fait de réordonner les colonnes de $\mathcal{S}_{:, > (i)}$ n'affecte pas les probabilités de $s_t^{(l)}$ conditionnellement à p_l pour tout $l \in \mathbf{L}_i$. En faisant un changement de variable $q = p/p_{(i)}$, on a :

$$\begin{aligned} \mathbb{P}(\mathcal{S}_{:, > (i)} = 0 \mid p_{(i)}) &= \lim_{N \rightarrow \infty} \left(\int_0^1 (1 - q p_{(i)})^T \frac{\alpha}{N} q^{\frac{\alpha}{N}-1} dq \right)^{N-i}, \\ &= \lim_{N \rightarrow \infty} \left(\int_0^1 (1 - q + q(1 - p_{(i)}))^T \frac{\alpha}{N} q^{\frac{\alpha}{N}-1} dq \right)^{N-i}, \\ &= \lim_{N \rightarrow \infty} \left(\int_0^1 \sum_{t=0}^T \binom{T}{t} (1 - q)^{T-t} (q(1 - p_{(i)}))^t \frac{\alpha}{N} q^{\frac{\alpha}{N}-1} dq \right)^{N-i}, \end{aligned}$$

$$= \lim_{N \rightarrow \infty} \left(\sum_{t=0}^T \binom{T}{t} \frac{\alpha}{N} (1-p_{(i)})^t \int_0^1 (1-q)^{T-t} q^{\frac{\alpha}{N}-1+t} dq \right)^{N-i},$$

où la troisième égalité provient de l'expansion de la série binomiale. Ceci mène à

$$\begin{aligned} \mathbb{P}(\mathcal{S}_{:, > (i)} = 0 \mid p_{(i)}) &= \lim_{N \rightarrow \infty} \left(\sum_{t=0}^T \binom{T}{t} \frac{\alpha}{N} (1-p_{(i)})^t \frac{\Gamma(T-t+1)\Gamma(\frac{\alpha}{N}+t)}{\Gamma(T+\frac{\alpha}{N}+1)} \right)^{N-i}, \\ &= \lim_{N \rightarrow \infty} \left(\sum_{t=0}^T \frac{T!}{t!(T-t)!} \frac{\alpha}{N} (1-p_{(i)})^t \frac{(T-t)!\Gamma(\frac{\alpha}{N}+t)}{\Gamma(T+\frac{\alpha}{N}+1)} \right)^{N-i}, \\ &= \lim_{N \rightarrow \infty} \left(\frac{T!\Gamma(\frac{\alpha}{N}+1)}{\Gamma(T+\frac{\alpha}{N}+1)} + \sum_{t=1}^T \frac{T!}{t!} \frac{\alpha}{N} (1-p_{(i)})^t \frac{\Gamma(\frac{\alpha}{N}+t)}{\Gamma(T+\frac{\alpha}{N}+1)} \right)^{N-i}, \\ &= \lim_{N \rightarrow \infty} \left[\frac{T!}{\prod_{t=1}^T (\frac{\alpha}{N}+t)} \left(1 + \frac{\alpha}{N} \sum_{t=1}^T (1-p_{(i)})^t \frac{\prod_{k=1}^{t-1} (\frac{\alpha}{N}+k)}{t!} \right) \right]^{N-i}. \end{aligned}$$

Or, d'après l'équation (C.2.3),

$$\lim_{N \rightarrow \infty} \left(\frac{T!}{\prod_{t=1}^T (\frac{\alpha}{N}+t)} \right)^{N-i} = \lim_{N \rightarrow \infty} \left(\frac{T!}{\prod_{t=1}^T (\frac{\alpha}{N}+t)} \right)^N = \exp \left(-\alpha \sum_{t=1}^T \frac{1}{t} \right).$$

De plus,

$$\lim_{N \rightarrow \infty} \sum_{t=1}^T (1-p_{(i)})^t \frac{\prod_{k=1}^{t-1} (\frac{\alpha}{N}+k)}{t!} = \sum_{t=1}^T \frac{(1-p_{(i)})^t}{t},$$

et

$$\lim_{N \rightarrow \infty} \frac{\alpha}{N} \sum_{t=1}^T (1-p_{(i)})^t \frac{\prod_{k=1}^{t-1} (\frac{\alpha}{N}+k)}{t!} = 0.$$

Puisque

$$\lim_{N \rightarrow \infty} \left(1 + \frac{X}{N} \right)^N = \exp(X),$$

$$\begin{aligned} \lim_{N \rightarrow \infty} \left(1 + \frac{\alpha}{N} \sum_{t=1}^T (1-p_{(i)})^t \frac{\prod_{k=1}^{t-1} (\frac{\alpha}{N}+k)}{t!} \right)^{N-i} &= \lim_{N \rightarrow \infty} \left(1 + \frac{\alpha}{N} \sum_{t=1}^T (1-p_{(i)})^t \frac{\prod_{k=1}^{t-1} (\frac{\alpha}{N}+k)}{t!} \right)^N, \\ &= \lim_{N \rightarrow \infty} \exp \left\{ \alpha \sum_{t=1}^T (1-p_{(i)})^t \frac{\prod_{k=1}^{t-1} (\frac{\alpha}{N}+k)}{t!} \right\}, \\ &= \exp \left(\alpha \sum_{t=1}^T \frac{(1-p_{(i)})^t}{t} \right). \end{aligned}$$

On obtient alors

$$\mathbb{P}(\mathcal{S}_{:, > (i)} = 0 \mid p_{(i)}) = \exp \left(-\alpha \sum_{t=1}^T \frac{1}{t} + \alpha \sum_{t=1}^T \frac{(1-p_{(i)})^t}{t} \right). \quad (\text{C.3.4})$$

Par la suite, pour obtenir la loi *a posteriori* de $p_{(N^*)}$, notons qu'il faut prendre en compte le fait que les colonnes d'indice supérieur à N^* sont nulles ($\mathcal{S}_{:, > N^*} = 0$), mais aussi que la colonne d'indice N^* est nulle

$(\mathcal{S}_{:,N^*} = 0)$. Ainsi, la probabilité à posteriori recherchée est obtenue par le produit des équations (C.3.2), (C.3.3) et (C.3.4), c'est-à-dire

$$\begin{aligned} \mathbb{P}(p_{(N^*)} \mid p_{(N^*-1)}, \mathcal{S}_{:, \geq N^*} = 0) &\propto \mathbb{P}(\mathcal{S}_{:, > N^*} = 0 \mid p_{(N^*)}) \mathbb{P}(\mathcal{S}_{:, N^*} = 0 \mid p_{(N^*)}) \mathbb{P}(p_{(N^*)} \mid p_{(N^*-1)}), \\ &\propto \exp\left(\alpha \sum_{t=1}^T \frac{(1 - p_{(N^*)})^t}{t}\right) (1 - p_{(N^*)})^T \left(p_{(N^*)}^{\alpha-1} \mathbf{1}_{\{0 \leq p_{(N^*)} \leq p_{(N^*-1)}\}}\right). \end{aligned}$$

□

C.4. Les autres distributions *a posteriori* de la section 4.3.2

En supposant qu'il y ait N_+ composantes actives et conditionnellement à \mathcal{S} et $c_{1:N_+}$, l'équation (4.3.3) stipule que $r_t \sim \mathcal{N}(0, V_t)$, où $V_t = \prod_{i=1}^{N_+} c_i^{s_t^{(i)}}$. En posant $s_t = \sum_{i=1}^{N_+} s_t^{(i)}$, $s^{(i)} = \sum_{t=1}^T s_t^{(i)}$ et $s = \sum_{t=1}^T s_t$, on a :

$$\mathbb{P}(r_t \mid \mathcal{S}, c_{1:N_+}) = \frac{1}{\sqrt{2\pi V_t}} \exp\left(-\frac{r_t^2}{2V_t}\right) = \frac{1}{\sqrt{2\pi \prod_{i=1}^{N_+} c_i^{s_t^{(i)}}}} \exp\left(-\frac{r_t^2}{2 \prod_{i=1}^{N_+} c_i^{s_t^{(i)}}}\right).$$

Ainsi,

$$\begin{aligned} \mathbb{P}(r_{1:T} \mid \mathcal{S}, c_{1:N_+}) &= \prod_{t=1}^T \mathbb{P}(r_t \mid \mathcal{S}, c_{1:N_+}) = \prod_{t=1}^T \left[\frac{1}{\sqrt{2\pi \prod_{i=1}^{N_+} c_i^{s_t^{(i)}}}} \right] \exp\left(-\sum_{t=1}^T \frac{r_t^2}{2 \prod_{i=1}^{N_+} c_i^{s_t^{(i)}}}\right), \\ &= (2\pi)^{-\frac{T}{2}} \prod_{i=1}^{N_+} \left(c_i^{-\frac{1}{2}s^{(i)}}\right) \exp\left(-\frac{1}{2} \sum_{t=1}^T \frac{r_t^2}{\prod_{i=1}^{N_+} c_i^{s_t^{(i)}}}\right). \end{aligned} \quad (\text{C.4.1})$$

$p_{(1:N_+)}$: Posons $p_{(0)} = 1$ et $p_{(N_++1)} \leq 1$. Ainsi, pour tout $i = 1, \dots, N_+$, la distribution *a posteriori* des probabilités $p_{(i)}$ conditionnellement aux autres variables est donnée par :

$$\begin{aligned} \mathbb{P}(p_{(i)} \mid rest) &\propto \mathbb{P}(p_{(i)} \mid \mathcal{S}, p_{1:(i-1)}, p_{(i+1):N_+}), \\ &\propto \mathbb{P}(\mathcal{S} \mid p_{(1:N_+)}) \mathbb{P}(p_{(i)} \mid p_{(i+1)}, p_{(i+1)}). \end{aligned}$$

En utilisant la loi conditionnelle de \mathcal{S} à la proposition 4.2.1 et la loi *a priori* de $p_{(i)}$ définie à la section 4.2.3, on a :

$$\begin{aligned} \mathbb{P}(p_{(i)} \mid rest) &\propto \left(\prod_{i=1}^{N_+} p_i^{n_{01}^{(i)} + n_{10}^{(i)}} (1 - p_i)^{n_{00}^{(i)} + n_{11}^{(i)}}\right) \left(\alpha p_{(1)}^{\alpha-1} \left[\prod_{i=1}^{N_++1} \alpha p_{(i)}^{\alpha-1} p_{(i-1)}^{-\alpha}\right] \mathbf{1}_{\{p_{(N_++1)} < p_{(N_+)} < \dots < p_{(1)} < p_{(0)}\}}\right), \\ &\propto p_{(i)}^{n_{01}^{(i)} + n_{10}^{(i)} - 1} (1 - p_{(i)})^{n_{00}^{(i)} + n_{11}^{(i)}} \mathbf{1}_{\{p_{(i+1)} < p_{(i)} < p_{(i-1)}\}}, \\ p_{(i)} \mid rest &\sim \text{Beta}\left(n_{01}^{(i)} + n_{10}^{(i)}, n_{00}^{(i)} + n_{11}^{(i)} + 1\right) \mathbf{1}_{\{p_{(i+1)} < p_{(i)} < p_{(i-1)}\}}. \end{aligned}$$

$c_{1:N_+}$: Pour tout $i = 1, \dots, N_+$, la distribution *a posteriori* des volatilités c_i conditionnellement aux autres variables est donnée par :

$$\begin{aligned} \mathbb{P}(c_i \mid rest) &\propto \mathbb{P}(c_i \mid \mathcal{S}, c_{1:i-1}, c_{i+1:N_+}, r_{1:T}), \\ &\propto \mathbb{P}(r_{1:T} \mid \mathcal{S}, c_{1:N_+}) \mathbb{P}(c_i). \end{aligned}$$

En utilisant l'équation (C.4.1) et la distribution à priori de c_i , on a :

$$\begin{aligned}
\mathbb{P}(c_i | rest) &\propto \left[(2\pi)^{-\frac{T}{2}} \prod_{j=1}^{N_+} \left(c_j^{-\frac{1}{2}s^{(j)}} \right) \exp \left(-\frac{1}{2} \sum_{t=1}^T \frac{r_t^2}{\prod_{j=1}^{N_+} c_j^{s_t^{(j)}}} \right) \right] \left[c_i^{-a_c-1} \exp \left(-\frac{b_c}{c_i} \right) \right], \\
&\propto c_i^{-\frac{1}{2}s^{(i)}} \exp \left(-\frac{1}{2c_i} \sum_{\substack{t=1 \\ s_t^{(i)}=1}}^T \frac{r_t^2}{\prod_{\substack{j=1 \\ j \neq i}}^{N_+} c_j^{s_t^{(j)}}} - \frac{1}{2} \sum_{\substack{t=1 \\ s_t^{(i)}=0}}^T \frac{r_t^2}{\prod_{\substack{j=1 \\ j \neq i}}^{N_+} c_j^{s_t^{(j)}}} \right) c_i^{-a_c-1} \exp \left(-\frac{b_c}{c_i} \right), \\
&\propto c_i^{-\frac{1}{2}s^{(i)}-a_c-1} \exp \left[-\frac{1}{c_i} \left(b_c + \frac{1}{2} \sum_{\substack{t=1 \\ s_t^{(i)}=1}}^T \frac{r_t^2}{\prod_{\substack{j=1 \\ j \neq i}}^{N_+} c_j^{s_t^{(j)}}} \right) \right], \\
c_i | rest &\sim \text{Inv-Gamma} \left(a_c + \frac{s^{(i)}}{2}, b_c + \frac{1}{2} \sum_{t=1}^T \frac{r_t^2 s_t^{(i)}}{\prod_{\substack{j=1 \\ j \neq i}}^{N_+} c_j^{s_t^{(j)}}} \right).
\end{aligned}$$

α : La distribution *a posteriori* de α conditionnellement aux autres variables est donnée par :

$$\begin{aligned}
\mathbb{P}(\alpha | rest) &\propto \mathbb{P}(\alpha | p_{(1:N_++1)}), \\
&\propto \mathbb{P}(p_{(1:N_++1)} | \alpha) \mathbb{P}(\alpha),
\end{aligned}$$

En utilisant la distribution de $p_{(1:N_++1)}$ définie à la section 4.2.3 et la loi *a priori* de α ($\text{Gamma}(a_\alpha, b_\alpha)$), on a :

$$\begin{aligned}
\mathbb{P}(\alpha | rest) &\propto \left(\alpha p_{(1)}^{\alpha-1} \left[\prod_{i=1}^{N_++1} \alpha p_{(i)}^{\alpha-1} p_{(i-1)}^{-\alpha} \right] \mathbf{1}_{\{p_{(N_++1)} < p_{(N_+)} < \dots < p_{(1)} < p_{(0)}\}} \right) \left(\alpha^{a_\alpha} \exp(-\alpha b_\alpha) \right), \\
&\propto \alpha^{N_++1} p_{(N_++1)}^{\alpha-1} \alpha^{a_\alpha} \exp(-\alpha b_\alpha), \\
&\propto \alpha^{a_\alpha+N_++1} \exp(\alpha \log p_{(N_++1)}) \exp(-\alpha b_\alpha), \\
\alpha | rest &\sim \text{Gamma}(a_\alpha + N_+ + 2, b_\alpha - \log p_{(N_++1)}).
\end{aligned}$$

C.5. Éléments complémentaires sur les résultats d'estimations

Les figures C.1, C.2, C.3 et C.4 présentent respectivement les résultats de simulations du modèles iFHMV sur les séries de données des indices financiers (NASDAQ 100 et FTSE 100) et d'actions (GES et CS). Il s'agit d'une estimation par échantillonnage de Gibbs par blocs. L'exercice d'estimation consiste en deux exécutions : une première de 11 250 itérations (temps de chauffe) et une seconde de 3 750 itérations. La deuxième exécution montre une convergence du nombre de composantes inféré N_+ et de la fonction de vraisemblance.

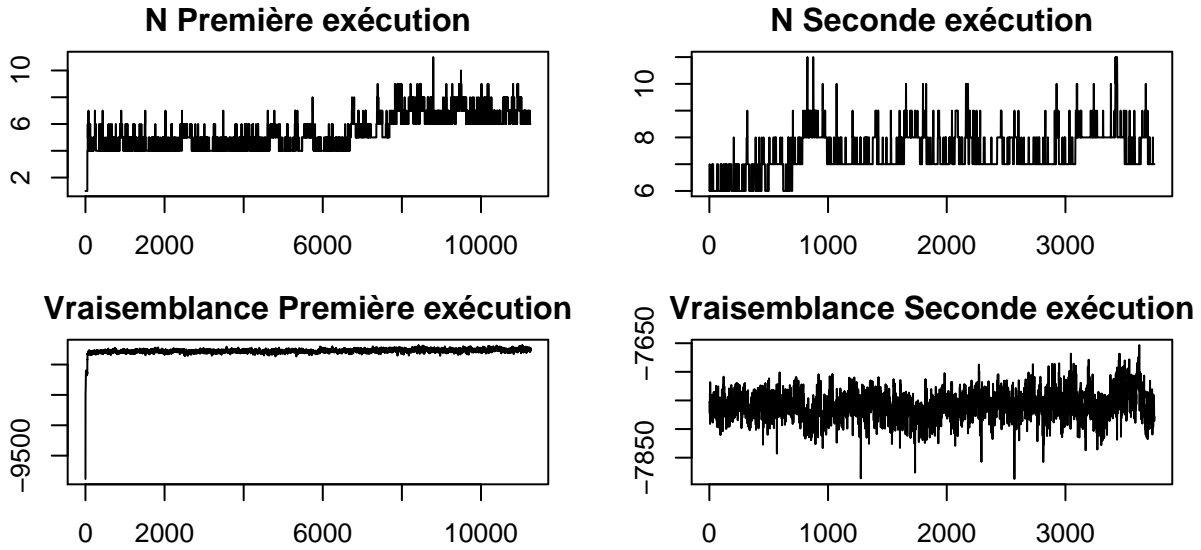


FIGURE C.1 – Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'indices financiers NASDAQ 100. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance.

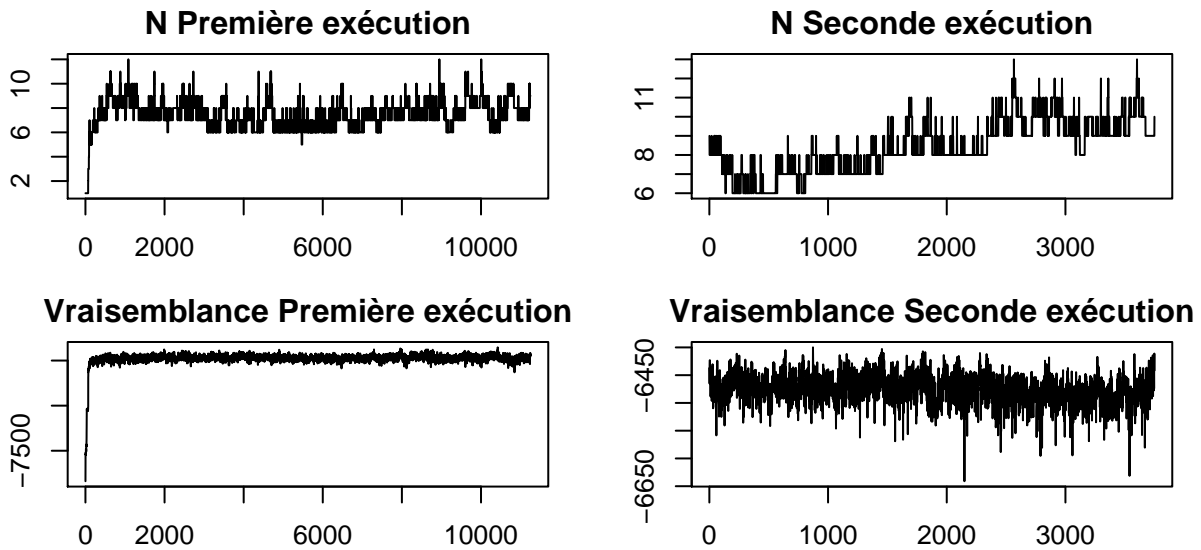


FIGURE C.2 – Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'indices financiers FTSE 100. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance.

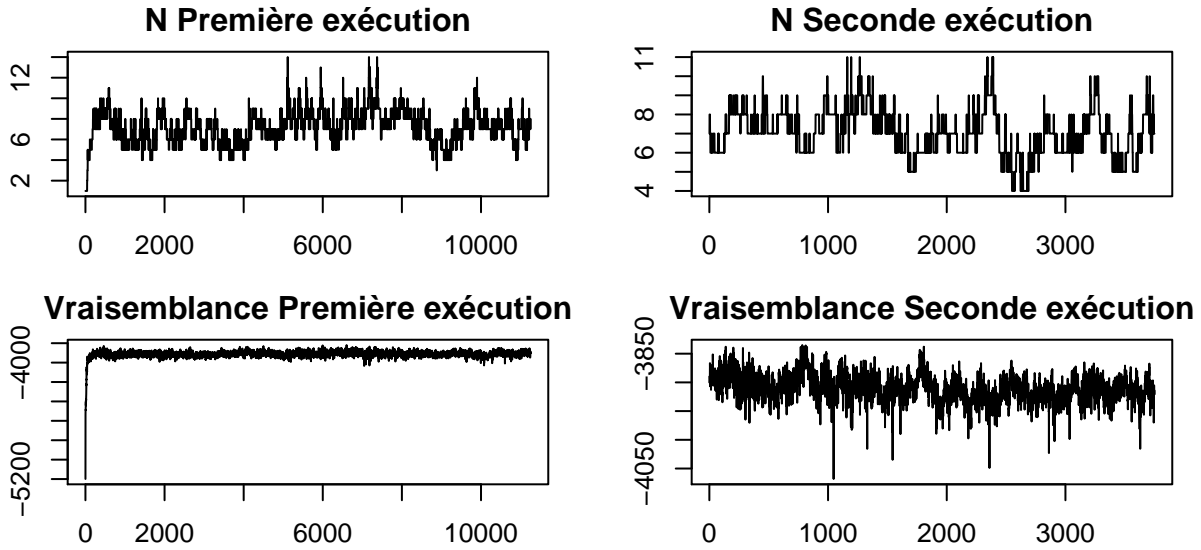


FIGURE C.3 – Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'action GES. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance.

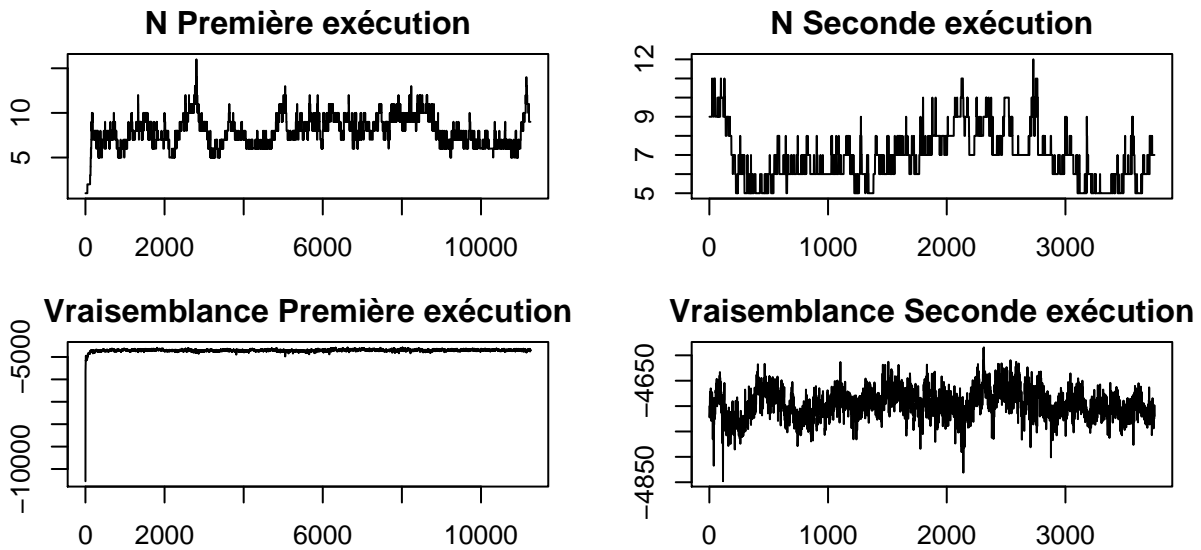


FIGURE C.4 – Résultats d'estimation par l'échantillonnage de Gibbs par blocs sur les données d'actions CS. Les deux colonnes concernent respectivement la première et la deuxième exécution. La première ligne présente le nombre de composantes N_+ inféré et la deuxième ligne présente la log-vraisemblance.

