Robustesse de la stratégie de trading optimale

Ahmed Bel Hadj Ayed

NNT : 2016SACLC033

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS-SACLAY,

préparée á CentraleSupélec.

ÉCOLE DOCTORALE N°573

Approches interdisciplinaires : Fondements, Applications et Innovation.

Mathématiques appliquées

Présentée par

Ahmed BEL HADJ AYED

Robustesse de la stratégie de trading optimale
Robustness of the optimal trading strategy

Thèse soutenue à Châtenay-Malabry, le 12/04/2016.

Composition du jury:

Denis TALAY	INRIA	*Président*
Frédéric ABERGEL	CentraleSupélec	*Directeur de Thèse*
Grégoire LOEPER	Monash University	*Co-directeur de Thèse*
Mathieu ROSENBAUM	Pierre et Marie Curie	*Rapporteur*
Huyên PHAM	Paris Diderot	*Rapporteur*
Damien CHALLET	CentraleSupélec	*Examinateur*

Laboratoire de Mathématiques et Informatique pour la Complexité et les Systèmes, CentraleSupélec.

ÉCOLE DOCTORALE
INTERFACES
Approches Interdisciplinaires : Fondements, Applications et Innovation

---

Titre: Robustesse de la stratégie de trading optimale

Mots-clés: trading, robustesse, calcul stochastique, calibration, estimation de tendance, filtre de Kalman, Merton, analyse technique, finance quantitative, ratio de Sharpe.

Résumé: L'objectif principal de cette thèse est d'apporter de nouveaux résultats théoriques concernant la performance d'investissements basés sur des modèles stochastiques. Pour ce faire, nous considérons la stratégie optimale d'investissement dans le cadre d'un modèle d'actif risqué à volatilité constante et dont la tendance est un processus caché d'Ornstein Uhlenbeck. Dans le premier chapitre, nous présentons le contexte et les objectifs de cette étude. Nous présentons, également, les différentes méthodes utilisées, ainsi que les principaux résultats obtenus. Dans le second chapitre, nous nous intéressons à la faisabilité de la calibration de la tendance. Nous répondons à cette question avec des résultats analytiques et des simulations numériques. Nous clôturons ce chapitre en quantifiant également l'impact d'une erreur de calibration sur l'estimation de la tendance et nous exploitons les résultats pour détecter son signe. Dans le troisième chapitre, nous supposons que l'agent est capable de bien calibrer la tendance et nous étudions l'impact qu'a la non-observabilité de la tendance sur la performance de la stratégie optimale. Pour cela, nous considérons le cas d'une utilité logarithmique et d'une tendance observée ou non. Dans chacun des deux cas, nous explicitons la limite asymptotique de l'espérance et la variance du rendement logarithmique en fonction du ratio signal-sur-bruit et de la vitesse de retour à la moyenne de la tendance. Nous concluons cette étude en montrant que le ratio de Sharpe asymptotique de la stratégie optimale avec observations partielles ne peut dépasser $\frac{2}{3^{3/2}} * 100\%$ du ratio de Sharpe asymptotique de la stratégie optimale avec informations complètes. Le quatrième chapitre étudie la robustesse de la stratégie optimale avec une erreur de calibration et compare sa performance à une stratégie d'analyse technique. Pour y parvenir, nous caractérisons, de façon analytique, l'espérance asymptotique du rendement logarithmique de chacune de ces deux stratégies. Nous montrons, grâce à nos résultats théoriques et à des simulations numériques, qu'une stratégie d'analyse technique est plus robuste que la stratégie optimale mal calibrée.

ÉCOLE DOCTORALE
INTERFACES
Approches Interdisciplinaires : Fondements, Applications et Innovation

Title: Robustness of the optimal trading strategy

Keywords: trading, robustness, stochastic calculus, trend filtering, inference of hidden Markov models, Kalman filtering, Merton, technical analysis, quantitative finance, Sharpe ratio.

Abstract: The aim of this thesis is to study the robustness of the optimal trading strategy. The setting we consider is that of a stochastic asset price model where the trend follows an unobservable Ornstein-Uhlenbeck process. In the first chapter, the background and the objectives of this study are presented along with the different methods used and the main results obtained. The question addressed in the second chapter is the estimation of the trend of a financial asset, and the impact of misspecification. Motivated by the use of Kalman filtering as a forecasting tool, we study the problem of parameters estimation, and measure the effect of parameters misspecification. Numerical examples illustrate the difficulty of trend forecasting in financial time series.The question addressed in the third chapter is the performance of the optimal strategy, and the impact of partial information. We focus on the optimal strategy with a logarithmic utility function under full or partial information. For both cases, we provide the asymptotic expectation and variance of the logarithmic return as functions of the signal-to-noise ratio and of the trend mean reversion speed. Finally, we compare the asymptotic Sharpe ratios of these strategies in order to quantify the loss of performance due to partial information. The aim of the fourth chapter is to compare the performances of the optimal strategy under parameters mis-specification and of a technical analysis trading strategy. For both strategies, we provide the asymptotic expectation of the logarithmic return as functions of the model parameters. Finally, numerical examples find that an investment strategy using the cross moving averages rule is more robust than the optimal strategy under parameters misspecification.

Remerciements

En premier lieu, je tiens à remercier mes deux directeurs de thèse, Frédéric Abergel et Grégoire Loeper, pour la confiance qu'ils m'ont accordée en acceptant d'encadrer ce travail doctoral, pour leur multiples conseils et pour toutes les heures qu'ils ont consacrées à diriger cette recherche. C'est grâce à eux que cette thèse fut une expérience très enrichissante autant scientifiquement que humainement.
Je remercie chaleureusement Mathieu Rosenbaum et Huyên Pham qui ont aimablement accepté d'être rapporteurs de cette thèse. Je remercie aussi, très chaleureusement, Damien Challet et Denis Talay d'avoir aimablement accepté d'honorer ma soutenance.
Je remercie une nouvelle fois Grégoire Loeper de m'avoir accueilli au sein de BNP Paribas. Au cours de cette très agréable expérience, j'ai énormément appris de sa vaste connaissance des marchés financiers. J'ai particulièrement apprécié son esprit de partage des connaissances. Je remercie également Jean-Philippe Lemor et Jean-Jacques Rabeyrin de m'avoir accordé le temps nécessaire pour rédiger mes papiers et mon manuscrit de thèse. Je souhaite remercier chaleureusement mes collègues de BNP Paribas avec lesquels j'ai eu le plaisir de travailler. Les discussions que j'ai pu avoir avec eux m'ont beaucoup aidé dans ce projet.
Je tiens à remercier tous les membres et ex-membres de l'équipe Fiquant que j'ai côtoyé tout au long de ma thèse. Je remercie en particulier, Sofiene El Aoud et Damien Challet pour les échanges que l'on a eu et pour le travail que l'on a fait ensemble. Je remercie Aymen Jedidi et Marouane Anane pour leurs conseils. Je remercie Dalia Ibrahim de sa relecture de mes différents papiers ainsi que de ses remarques judicieuses. Je remercie également Sylvie Dervin pour son soutien.
Ce travail est dédié à ma mère qui m'a toujours poussé et soutenu tout au long de mes études, à mon père, à mes sœurs, à ma femme et à ma fille.

Contents

Chapter 1

Contexte, Méthodes et Résultats (In French)

1.1 Contexte et objectifs

Selon l'hypothèse de l'efficience des marchés, les rendements des actifs sont, théoriquement, bien décrits par une marche aléatoire. Cette fondation théorique, introduite dans Bachelier (1900), affirme que dans le marché, le prix d'un actif représente l'équilibre entre l'offre et la demande et est donc égal à sa « valeur fondamentale ». Fama (1969) formalisa cette définition en affirmant que, dans un marché efficient, le prix d'un actif suit un processus aléatoire, non prévisibles car sa variation dépend à chaque fois d'une nouvelle information elle-même non prévisible. Si c'est le cas, cela veut dire qu'aucune stratégie d'investissement ne pourrait être rentable en moyenne.
Cependant, tout le monde n'est pas convaincu de cette impossibilité à prédire l'avenir et une partie non négligeable de l'industrie financière travaille sur cette problématique. En effet, prenons par exemple les "Commodity Trading Advisors" qui sont des fonds d'investissement sur des futures et des contrats à termes. Ces structures prennent une grande partie de leurs profits grâce à des stratégies spéculatives dites "suiveuses de tendances" (voir Lempérière *et al.* (2014)). Un autre exemple très connu est celui de Warren Buffet, qui pendant plus de 40 ans, a toujours surpassé les indices de référence tels que le Dow Jones. D'un point de vue empirique, il existe une grosse littérature sur le sujet et certains articles ont montré que les prix d'actifs présentent des anomalies (voir Schwert (2003)) et que les performances passées tendent à justifier l'existence d'une tendance (Lempérière *et al.* (2014) ou encore Anane & Abergel (2015)).
Du côté des mathématiques, un grand nombre de chercheurs se sont intéressés à la notion d'investissement optimal, en supposant de façon sous-jacente qu'il existait une tendance. Le premier est Markovitz (1952). Sa théorie décrit le comportement que devrait suivre un investisseur pour con-

struire un portefeuille dans un univers incertain (voir Clauss (2011)). Sa formalisation la plus accomplie est le modèle d'évaluation des actifs financiers ou MEDAF (voir Treynor (1961) ou Sharpe (1964) pour plus de détails). Ces théories concluent que l'investisseur doit chercher à optimiser le couple rentabilité-risque d'un portefeuille donné. Il s'agit donc d'une approche de type moyenne-variance. En 1969, Robert C. Merton introduisit sa théorie sur l'investissement optimal (voir Merton (1990) pour plus de détails). Le profil de risque de l'investisseur est caractérisé par une fonction d'utilité que l'on souhaite maximiser. Celui-ci travaille sur le modèle d'actif classique de Black, Scholes et Merton (voir Black & Scholes (1973)) et utilise une méthode mathématique nommée le contrôle stochastique pour la résolution de ce problème. Peu importe les généralisations de ce problème initial (voir J.C. Cox (1989), JJ.C. Cox (1985), D. Duffie (1989), H. He (1991), I. Karatzas (1987), D.L. Ocone (1991), ou encoreLakner (1995)), elles font toutes intervenir une tendance connue ou au moins avec une dynamique connue. C'est d'ailleurs cette hypothèse d'une tendance correctement estimée qui rend difficile la mise en pratique de ce genre de stratégie. Pour y arriver, les intervenants du marché utilisent la méthode du backtest sur des données passées pour calibrer les différents paramètres d'un modèle au risque d'introduire de la sur-optimisation (voir Challet & Bel Hadj Ayed (2015) où est rappelé comment effectuer un backtest).
L'objectif de ce travail est d'apporter des résultats théoriques comblant ce fossé entre la littérature et l'industrie. Pour y parvenir, nous considérons dans cette thèse la stratégie optimale de trading avec une utilité logarithmique dans le cadre d'un modèle d'actif risqué à volatilité constante et dont la tendance est un processus caché d'Ornstein Uhlenbeck.
Assumant donc dans ce manuscrit que la tendance existe, nous nous intéressons dans un premier temps à la faisabilité de la calibration, l'estimation et la détection de la tendance.
Admettant ensuite que nous sommes capables de connaitre la dynamique de la tendance, la second partie de cette thèse quantifie l'impact de la non-observabilité de celle-ci sur la performance de l'investissement optimal.
Finalement, au vu des résultats obtenus dans la première partie, le dernier chapitre s'intéresse à cette stratégie lorsque le modèle de tendance est mal calibré, et compare la stabilité de sa performance à celle d'une stratégie d'analyse technique.

Dans la suite de ce chapitre, nous présentons les différentes parties de ce travail en résumant les méthodes et les résultats.

1.2 Calibration, estimation et détection de la tendance avec les prix historiques

Comme expliqué ci-dessus, une grande partie de l'industrie financière admet l'existence de la tendance. De plus, une grande majorité des stratégies d'investissements fondées sur des modèles mathématiques suppose l'existence d'une tendance sous-jacente (voir Markovitz (1952) ou Merton (1990) par exemple).
Il est alors naturel de se demander si la calibration ou l'estimation d'un tel processus est possible. Même dans un cadre théorique simplifié, il s'agit d'un problème statistique très difficile en raison du bruit élevé sur les observations. Supposons par exemple le modèle suivant avec une tendance constante : $\frac{dS_t}{S_t} = \mu dt + \sigma_S dW_t^S$. Dans ce cas, la meilleure estimation de la tendance à un instant T est donnée par $\widehat{\mu}_t = \frac{1}{T}\int_0^T \frac{dS_u}{S_u}$. Le test de Student rejettera l'hypothèse $\mu = 0$ à un niveau d'erreur de 5% si $|\widehat{\mu}_T| > \frac{1.96\sigma_S}{\sqrt{T}}$. Si par exemple $\sigma_S = 30\%$, une estimée $\widehat{\mu}_T = 1\%$ deviendra statistiquement viable après $T > 3457$ années...
L'objectif de ce chapitre est alors de valider la faisabilité de la calibration de la tendance modélisée pas un processus aléatoire non-observé avec un comportement de type retour à la moyenne. Utilisant une approche Bayésienne ou de type Maximum de vraisemblance (voir Leroux (1992),O. Cappé & Moulines (2005), Benmiloud & Pieczynski (1995), Casarin & Marin (2007) ou Dahia (2005)), des méthodes d'inférences sur des processus cachés ont été appliquées au séries temporelles financières, mais la plupart de ces études se sont concentrées sur l'estimation d'une volatilité stochastique (voir Jacquier *et al.* (1994),Eraker (1998),Kim *et al.* (1998) ou Chib *et al.* (2002)).
Plusieurs auteurs ont étudié des processus de tendances cachés et utilisent alors des méthodes de filtrage pour estimer la tendance (voir Lakner (1998), Pham & Quenez (2001),Laskry & Lions (1999) ou Brendle (2006a)). La plupart de ces filtres sont basés sur un modèle paramétrique pour la tendance, et donc leur utilisation en pratique est compliquée car ces méthodes sont alors confrontées au problème de l'inférence des paramètres. C'est donc pour combler ce fossé entre la théorie et la pratique que nous nous sommes intéressés à la faisabilité de ce problème, et donc à la faisabilité de la mise en pratique de stratégies classiques de trading.

1.2.1 Modèle et filtrage linéaire

Considérons un marché financier défini sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$, où $\mathbf{F} = \{\mathcal{F}_t, t \geqslant 0\}$ est la filtration naturelle engendrée par l'association de deux processus Browniens décorrélés (W^S, W^μ), et $\mathbb{P}$ la mesure de probabilité historique. Dans ce chapitre, et aussi dans toute cette thèse (sauf indication contraire), nous considérons que l'actif risqué est décrit par le modèle suivant :

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_S dW_t^S, \tag{1.1}$$

$$d\mu_t = -\lambda \mu_t dt + \sigma_\mu dW_t^\mu, \tag{1.2}$$

où la tendance initiale $\mu_0 = 0$. Pour éviter les problèmes aux bords, on suppose également que $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$. Le paramètre λ est nommé la vitesse de retour à la moyenne de la tendance. En effet, ce paramètre peut être interprété comme une force ramenant la tendance vers zéro. Notons également $\mathbf{F}^S = \left\{\mathcal{F}_t^S\right\}$ la filtration engendrée par S. Il faut remarquer que seuls les processus adaptés à la filtration $\mathbf{F}^S$ sont observables. Cela veut dire que les agents sur ce marché n'observent pas la tendance. En pratique, les observations des prix sont discrètes sur le marché. Si l'on note δ la fréquence d'échantillonnage de notre système, et que l'on représente l'instant $t_k = k\delta$ par l'indice k, la version en temps discret de ce modèle peut alors s'écrire :

$$y_{k+1} = \frac{S_{k+1} - S_k}{\delta S_k} = \mu_{k+1} + u_{k+1}, \tag{1.3}$$

$$\mu_{k+1} = e^{-\lambda\delta}\mu_k + v_k, \tag{1.4}$$

où $u_k \sim \mathcal{N}\left(0, \frac{\sigma_S^2}{\delta}\right)$ et $v_k \sim \mathcal{N}\left(0, \frac{\sigma_\mu^2}{2\lambda}\left(1 - e^{-2\lambda\delta}\right)\right)$.

Si l'on suppose que tous les paramètres de ce modèle sont connus, il est possible de fournir une estimation optimale de la tendance. En effet, le système (1.3)-(1.4) correspond alors à un modèle espace-état linéaire et Gaussien où les observations sont représentées par y et l'état du système par μ (voir Brockwell & Davis (2002a) pour plus de précisions). Dans ce cas, l'estimation optimale de la tendance μ est donnée par le filtrage de Kalman, qui permet d'obtenir l'espérance conditionnelle $E\left[\mu_t | \mathcal{F}_t^S\right]$. Vu que $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$, le système est alors contrôlable, observable et invariant dans le temps. Dans ce cas, il est bien connu que la variance du filtre de Kalman, représentant l'erreur d'estimation, converge vers une valeur constante (voir Kalman *et al.* (1962) pour plus de détails). On parle alors du filtre de Kalman en régime permanent ou stationnaire. En temps discret, le filtre de Kalman en régime permanent peut s'écrire :

$$\widehat{\mu}_{n+1} = K_\infty \sum_{i=0}^{\infty} e^{-\lambda\delta i} \left(1 - K_\infty\right)^i y_{n+1-i}, \tag{1.5}$$

où le gain stationnaire K_∞ dépend des paramètres du système. Ce filtre peut également être défini en temps continu. Dans ce cas, on peut écrire la proposition suivante :

Proposition 1. *Le filtre de Kalman en temps continu et en régime permanent $\widehat{\mu}$ est solution de l'équation stochastique différentielle suivante :*

$$d\widehat{\mu}_t = -\lambda\beta\left(\lambda,\sigma_\mu,\sigma_S\right)\widehat{\mu}_t dt + \lambda\left(\beta\left(\lambda,\sigma_\mu,\sigma_S\right)-1\right)\frac{dS_t}{S_t}, \tag{1.6}$$

où

$$\beta\left(\lambda,\sigma_\mu,\sigma_S\right) = \left(1+\frac{\sigma_\mu^2}{\lambda^2\sigma_S^2}\right)^{\frac{1}{2}}. \tag{1.7}$$

En pratique, même si l'on admet pouvoir estimer correctement la volatilité, les paramètres $\theta = (\lambda, \sigma_\mu)$ sont inconnus et doivent être estimés. Ce point est traité dans la partie suivante.

1.2.2 Inférence du modèle avec des observations discrètes

Dans cette partie, l'inférence des paramètres à partir d'observations en temps discret est traitée. Que ce soit par approche bayésienne ou par maximum de vraisemblance, l'estimation fait intervenir la fonction de vraisemblance. Pour le calcul de cette fonction, nous proposons deux méthodes. La première est un calcul direct. Vu que les observations $(y_1,\cdots,y_N)^T$, sachant les paramètres θ, forment un processus Gaussien, la vraisemblance des observations est alors caractérisée par la moyenne $M_{y_{1:N}|\theta}$ et la covariance $\Sigma_{y_{1:N}|\theta}$:

$$\begin{aligned} M_{y_{1:N}|\theta} &= 0, & (1.8)\\ \Sigma_{y_{1:N}|\theta} &= \Sigma_{\mu_{1:N}|\theta} + \Sigma_{u_{1:N}|\theta}, & (1.9)\end{aligned}$$

où $\Sigma_{u_{1:N}|\theta} = \frac{\sigma_S^2}{\delta} I_N$ et $\Sigma_{\mu_{1:N}|\theta} = \left(\mathbb{C}ov\left(\mu_t,\mu_s\right)\right)_{1\leq t,s\leq N}$. Vu que le drift μ est un processus d'Ornstein Uhlenbeck :

$$\mathbb{C}ov\left(\mu_t,\mu_s\right) = \frac{\sigma_\mu^2}{2\lambda}e^{-\lambda(s+t)}\left(e^{2\lambda s\wedge t}-1\right). \tag{1.10}$$

Ainsi la fonction de vraisemblance s'écrit :

$$f\left(y_1,...y_N|\theta\right) = \frac{1}{(2\pi)^{N/2}\sqrt{det\Sigma_{y_{1:N}|\theta}}}e^{\left(\frac{-1}{2}(y_1,...,y_N)\Sigma_{y_{1:N}|\theta}^{-1}(y_1,...,y_N)^T\right)}. \tag{1.11}$$

Lorsque le nombre d'observations N est très grand, il devient impossible de calculer directement l'inverse et le déterminant de la matrice de covariance $\Sigma_{y_{1:N}|\theta}$. Nous proposons alors une méthode récursive permettant d'effectuer ce calcul direct de vraisemblance.

Un seconde méthode consiste à calculer de façon récursive la vraisemblance en utilisant la décomposition suivante (voir Schweppe (1965) pour plus de détails) :

$$\begin{aligned} f\left(y_1,...y_N|\theta\right) &= f\left(y_N|y_1,...y_{N-1},\theta\right) f\left(y_1,...y_{N-1}|\theta\right). \\ &= \prod_{n=1}^{N} f\left(y_n|y_1,...y_{n-1},\theta\right), \end{aligned}$$

et nous pouvons utiliser la proposition ci-dessous pour expliciter les densités conditionnelles :

Proposition 2. *Le processus* $(y_n|y_1,...y_{n-1},\theta)$ *est Gaussien :*

$$(y_n|y_1,...y_{n-1},\theta) \quad \sim \quad \mathcal{N}\left(M_{y_{n|n-1}}, \mathbb{V}ar_{y_{n|n-1}}\right),$$

avec

$$\begin{aligned} M_{y_{n|n-1}} &= e^{-\lambda\delta}\hat{\mu}_{n-1/n-1}, \\ \mathbb{V}ar_{y_{n|n-1}} &= e^{-2\lambda\delta}\Gamma_{n-1/n-1} + \frac{\sigma_\mu^2}{2\lambda}\left(1-e^{-2\lambda\delta}\right) + \frac{\sigma_S^2}{\delta}. \end{aligned}$$

L'estimation a posteriori de la tendance $\hat{\mu}_{n-1/n-1}$ *et la variance a posteriori de ce filtre* $\Gamma_{n-1/n-1}$ *sont fournies par le filtrage de Kalman.*

Il est alors possible de mettre en pratique l'inférence des paramètres. Nous nous sommes alors intéressés aux comportements asymptotiques de ces estimateurs. C'est en utilisant le fait que le modèle en temps discret peut s'écrire comme un processus ARMA(1,1) (voir Genon-Catalot (2009) pour plus de détails), que nous avons pu affirmer que l'estimateur du maximum de vraisemblance et les estimateurs Bayésiens sont asymptotiquement normales. De plus, nous savons grâce à la littérature sur les processus ARMA Gaussiens (voir par exemple Brockwell & Davis (2002b), section 10.8), que si $\hat{\theta}_N$ est un estimateur non biaisé, on a alors :

$$\mathbb{C}ov_\theta\left(\hat{\theta}_N\right) \geqslant CRB\left(\theta\right),$$

où $CRB\left(\theta\right)$ est la borne de Cramer Rao et est donnée par $CRB\left(\theta\right) = I_N^{-1}\left(\theta\right)$, où $I_N\left(\theta\right)$ est la matrice d'information de Fisher :

$$\left(I_N\left(\theta\right)\right)_{i,j} = -\mathbb{E}\left[\frac{\partial^2 \log f\left(y_1,...y_N|\theta\right)}{\partial\theta_i\partial\theta_j}\right],$$

L'information de Fisher peut être obtenue de façon analytique grâce à ce théorème :

Theorem 1.2.1. *Considérons le modèle* (1.3)-(1.4). *Si* $(\lambda,\sigma_\mu,\sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$*, alors:*

$$I_1\left(\theta\right) = \left(\frac{1}{4\Pi}\int_{-\Pi}^{\Pi} f_\theta^{-2}\left(\omega\right)\frac{\partial f_\theta}{\partial\theta_i}\left(\omega\right)\frac{\partial f_\theta}{\partial\theta_j}\left(\omega\right)d\omega\right)_{1\leq i,j\leq 2},$$

où f_θ est la densité spectrale du processus (y_i):

$$f_\theta(\omega) = \frac{\frac{\sigma_\mu^2}{2\lambda}\left(1 - e^{-2\lambda\delta}\right) + \frac{\sigma_S^2}{\delta}\left(1 + e^{-2\lambda_\mu\delta}\right) - \frac{2e^{-\lambda_\mu\delta}\sigma_S^2}{\delta}\cos(\omega)}{1 + e^{-2\lambda_\mu\delta} - 2e^{-\lambda_\mu\delta}\cos(\omega)}.$$

Il est alors possible de calculer la Borne de Cramer Rao et de savoir combien de temps il faut observer le système pour atteindre une erreur cible sur un paramètre θ_i (caractérisée par un écart-type sur l'estimateur x_i). Par exemple pour une fréquence d'échantillonnage $\delta = 1/252$, donc à partir de rendements journaliers, il faut T_i^x années pour atteindre un écart-type de x_i sur l'estimation de θ_i :

$$T_i^x = \frac{\left(I_1^{-1}(\theta)\right)_{ii}}{252 * x_i^2}.$$

Nous avons alors calculé cette durée sur différents régimes de tendances en fixant une volatilité $\sigma_S = 30\%$. La plus courte période obtenue est supérieure à 29 années. Elle correspond au temps nécessaire pour avoir une erreur $x = 0.5$ sur le paramètre $\lambda_\mu = 1$ avec un écart type de la tendance égal à $\sigma_\mu (2\lambda_\mu)^{-1/2} \approx 63\%$. Pour cette configuration, après 30 années d'observations, l'écart-type de l'estimateur vaut 50% de la valeur du paramètre que l'on souhaite estimer. Si l'on veut atteindre un écart-type de 10%, il faut attendre 742 années. Même pour une tendance ayant un grand écart-type ($\approx 63\%$), la calibration de ce modèle en un temps raisonnable et avec une bonne précision est impossible.

1.2.3 Impact d'une erreur de calibration sur l'estimation et la détection de tendance en temps continu

Au vu des résultats de la partie précédente, nous nous sommes intéressés dans cette partie au filtre de Kalman en temps continu (et en régime permanent) avec une mauvaise calibration. Nous supposons alors que l'actif risqué est donné par le modèle (1.1)-(1.2) avec $\theta^* = \left(\sigma_\mu^*, \lambda^*\right)$, et nous considérons le cas d'un agent pensant que les vrais paramètres valent $\theta = (\sigma_\mu, \lambda)$. L'agent va alors implémenter le filtre suivant :

$$d\widehat{\mu}_t = -\lambda\beta\widehat{\mu}_t dt + \lambda(\beta - 1)\frac{dS_t}{S_t}, \tag{1.12}$$

où $\beta = \beta(\lambda, \sigma_\mu, \sigma_S)$. Après avoir caractérisé la loi de ce filtre, nous avons quantifié l'impact d'une mauvaise calibration en nous intéressant au processus de résidu que l'on définit comme étant la différence entre le filtre et le drift caché. Le théorème suivant donne la loi de ce résidu :

Theorem 1.2.2. *Considérons le modèle* (1.1)-(1.2) *avec* $\theta^* = \left(\sigma_\mu^*, \lambda^*\right)$ *et le filtre mal-spécifié défini dans l'équation (1.12). Le processus résidu* $\widehat{\mu} - \mu^*$ *est* *alors un processus Gaussien centré dont la variance a une limite stationnaire :*

$$\lim_{t\to\infty} \mathbb{V}ar\left[\widehat{\mu}_t - \mu_t^*\right] = \frac{\sigma_S^2}{2\beta}\left(\lambda\left(\beta-1\right)^2 + \lambda^*\left(\left(\beta^*\right)^2-1\right)\frac{\lambda^*\beta+\lambda}{\lambda\beta+\lambda^*}\right), \quad (1.13)$$

où $\beta = \beta\left(\lambda, \sigma_\mu, \sigma_S\right)$ *et* $\beta^* = \beta\left(\lambda^*, \sigma_\mu^*, \sigma_S\right)$ *comme défini dans l'équation (1.7).*
De plus, si le modèle est bien calibré $\left(\sigma_\mu, \lambda\right) = \left(\sigma_\mu^*, \lambda^*\right)$*, l'équation (1.13) devient :*

$$\lim_{t\to\infty} \mathbb{V}ar\left[\widehat{\mu}_t^* - \mu_t^*\right] = \lambda^*\sigma_S^2\left(\beta^*-1\right). \quad (1.14)$$

Nous montrons également que la variance relative asymptotique des résidus dans le cas bien-spécifié $\left(\lim_{t\to\infty}\frac{\mathbb{V}ar\left[\widehat{\mu}_t^*-\mu_t^*\right]}{\mathbb{V}ar\left[\mu_t^*\right]}\right)$ est une fonction croissante de λ^* et une fonction décroissante de σ_μ^*.
Avec ces résultats théoriques, nous avons pu étudier au travers de simulations numériques, l'impact d'une mauvaise calibration sur l'estimation de la tendance une volatilité $\sigma_S = 30\%$.
Par exemple, dans le cas d'un filtre de Kalman bien calibré, pour $\lambda^* = 1$ et $\sigma_\mu^* = 90\%$, l'écart-type des résidus ($\simeq 44\%$) est inférieur à l'écart-type de la tendance ($\simeq 64\%$). Pour des valeurs élevées de λ^* et des valeurs faibles de σ_μ^*, les deux quantités sont équivalentes. Nous retrouvons alors le fait que le filtrage est de meilleure qualité pour de faibles valeurs de λ^* et de grandes valeurs de σ_μ^*.
Considérons le cas où $\lambda^* = 1$ et $\sigma_\mu^* = 90\%$. Comme précisé ci-dessus, ce régime correspond à un écart-type de la tendance de $\sigma_\mu^*\left(2\lambda_\mu^*\right)^{-1/2} \approx 63\%$ et a un écart-type des résidus de 44% avec une bonne calibration. Si l'agent considère que $\lambda = 5$ et $\sigma_\mu = 10\%$, l'écart-type des résidus deviennent supérieurs à 60%.
Même lorsque le régime de la tendance semble avantageux, l'impact d'une mauvaise calibration est non négligeable.
Nous pouvons tout de même extraire une information exploitable de ce filtrage, qu'il soit bien ou mal-spécifié. En effet, nous nous sommes intéressés au problème de la détection d'une tendance positive (ou négative). En effet, nous avons caractérisé dans un premier temps la loi asymptotique du processus conditionnel $\left(\mu_t^*|\hat{\mu}_t = x\right)$:

Proposition 3. *Considérons le modèle* (1.1)-(1.2) *avec* $\theta^* = \left(\sigma_\mu^*, \lambda^*\right)$ *et le filtre mal-spécifié défini dans l'équation (1.12). Dans ce cas :*

$$\left(\mu_t^*|\hat{\mu}_t = x\right) \underset{t\to\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}\left(\mathbb{M}_{\mu^*|\hat{\mu}}^\infty, \mathbb{V}ar_{\mu^*|\hat{\mu}}^\infty\right), \quad (1.15)$$

avec:

$$\mathbb{M}^{\infty}_{\mu^*|\hat{\mu}} = \frac{\lambda^*_\mu \beta \left((\beta^*)^2 - 1 \right)}{(\beta - 1) \left(\lambda_\mu \beta + \lambda^*_\mu (\beta^*)^2 \right)} x, \tag{1.16}$$

$$\mathbb{V}ar^{\infty}_{\mu^*|\hat{\mu}} = \mathbb{V}ar^{\infty}_{\mu^*} \left(1 - \frac{\lambda^*_\mu \lambda_\mu \beta \left((\beta^*)^2 - 1 \right)}{\left(\lambda^*_\mu + \lambda_\mu \beta \right) \left(\lambda_\mu \beta + \lambda^*_\mu (\beta^*)^2 \right)} \right), \tag{1.17}$$

où $\mathbb{V}ar^{\infty}_{\mu^*} = \frac{(\sigma^*_\mu)^2}{2\lambda^*_\mu}$.
De plus, dans le cas bien spécifié $(\sigma_\mu, \lambda) = \left(\sigma^*_\mu, \lambda^* \right)$ *l'équation (1.15) devient:*

$$(\mu^*_t | \hat{\mu}^*_t = x) \underset{t \to \infty}{\overset{\mathcal{L}}{\to}} \mathcal{N} \left(x, \frac{2\mathbb{V}ar^{\infty}_{\mu^*}}{\beta^* + 1} \right), \tag{1.18}$$

où $\beta^* = \beta \left(\lambda^*, \sigma^*_\mu, \sigma_S \right)$ *(voir l'équation (1.7)).*

Ainsi, avec cette proposition, il est possible de déduire la prochaine proposition, qui nous donne la probabilité asymptotique d'avoir une tendance positive, sachant une estimé x positive :

Proposition 4. *Considérons le modèle* (1.1)-(1.2) *avec* $\theta^* = \left(\sigma^*_\mu, \lambda^* \right)$ *et le filtre mal-spécifié défini dans l'équation (1.12). Dans ce cas :*

$$\lim_{t \to \infty} \mathbb{P} \left(\mu^*_t > 0 | \hat{\mu}_t = x \right) = \mathbb{P}_\infty \left(\mu^* > 0 | \hat{\mu} = x \right), \tag{1.19}$$

où

$$\mathbb{P}_\infty \left(\mu^* > 0 | \hat{\mu} = x \right) = 1 - \Phi \left(\frac{-\mathbb{M}^{\infty}_{\mu^*|\hat{\mu}=x}}{\sqrt{\mathbb{V}ar^{\infty}_{\mu^*|\hat{\mu}=x}}} \right), \tag{1.20}$$

où $\mathbb{M}^{\infty}_{\mu^*|\hat{\mu}=x}$ *et* $\mathbb{V}ar^{\infty}_{\mu^*|\hat{\mu}=x}$ *sont définis dans les équations (1.16) et (1.17), et* Φ *est la fonction de répartition de la loi normale centrée réduite.*
De plus, si $x > 0$ *et que le modèle est bien calibré* $(\sigma_\mu, \lambda) = \left(\sigma^*_\mu, \lambda^* \right)$*, cette probabilité asymptotique devient une fonction croissante* σ^*_μ *et une fonction décroissante de* λ^*_μ*.*

Ainsi, grâce à la corrélation non nulle entre le processus de drift caché et le filtre (bien ou mal-spécifié), cette probabilité est toujours supérieure à 0.5 et peut donc être utilisée pour détecter une tendance positive.

1.2.4 Conclusion

Ce chapitre illustre la difficulté du filtrage de la tendance avec un modèle où le drift est un processus caché d'Ornstein Uhlenbeck. Ce modèle appartient

à la classe des systèmes espaces-états linéaires et Gaussiens. L'avantage de ce cadre est de présenter une méthode "temps réel" d'estimation : le filtrage de Kalman.
En pratique, ces paramètres sont inconnus et la calibration du modèle est essentielle. Le cadre régulier de cette modélisation permet d'obtenir, de façon analytique, la vraisemblance et le filtre de Kalman permet également de l'obtenir de façon récursive.
Malgré ces avantages, les résultats de cette analyse montrent que les estimateurs classiques ne sont pas adaptés à un si faible ratio signal-sur-bruit. L'horizon d'observations nécessaire pour avoir une précision suffisante est trop long. La convergence des estimateurs des paramètres n'est donc pas garantie et l'impact d'une mauvaise calibration est non-négligeable sur l'estimation de la tendance.
Malgré ces difficultés, la corrélation non nulle entre la tendance cachée et son estimateur (bien ou mal-spécifié) peut être utilisée pour la détection du signe de la tendance.

1.3 Impact de le non-observabilité de la tendance sur la performance de la stratégie optimale

La notion d'investissement optimal a été introduite par Merton en 1969 (voir Merton (1990) pour plus de détails). Partant d'un actif risqué modélisé par un mouvement Brownien géométrique, il dériva l'allocation optimale permettant de maximiser une fonction d'utilité future espérée. De nombreuses généralisations du problème initial sont possibles. L'une d'elle est de supposer que la tendance est un processus stochastique inobservé, ce qui revient donc à considérer un système avec informations partielles. C'est ce qui a été fait par Karatzas & Zhao (2001) qui ont considéré le cas d'une tendance constante cachée, par Brendle (2006b) qui supposa que la tendance est un processus d'Ornstein Uhlenbeck caché, et par Sass & Haussmann (2004) qui ont pris pour modèle de tendance une chaine de Markov cachée en temps continu.
Dans ce chapitre, nous considérons le cas développé par Brendle, une tendance modélisée par un processus d'Ornstein Uhlenbeck caché et nous considérons le problème de l'investissement optimal avec une fonction d'utilité logarithmique dans le cadre d'informations partielles ou complètes.
Le but de ce travail est dans un premier temps de caractériser la performance de ces deux stratégies (avec ou sans observabilité de la tendance) comme des fonctions du ratio signal-sur-bruit et de la vitesse de retour à la moyenne de la tendance. La finalité de l'étude est de caractériser l'impact de la non-observabilité de la tendance sur le performance de la stratégie optimale d'investissement.
La perte sur la fonction d'utilité espérée en raison d'observations partielles a

déjà été abordée dans Karatzas & Zhao (2001), dans Brendle (2006b) et dans Rieder & Bauerle (2005). Dans ce chapitre, la performance d'un portefeuille P est évaluée à l'aide du ratio de Sharpe annualisé (voir Sharpe (1966)) sur les rendements logarithmiques :

$$\mathrm{SR}_T = \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{\sqrt{T\ \mathbb{V}\mathrm{ar}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}}. \tag{1.21}$$

Le principe de cet indicateur est de mesurer le rendement logarithmique espéré par unité de risque. C'est une des mesures les plus utilisées dans l'industrie financière.

1.3.1 Préliminaires

Dans ce chapitre, nous considérons le modèle (1.1)-(1.2). Comme expliqué dans le premier chapitre, la tendance μ n'est pas observable par l'agent. Partant de ce cadre, il est tout de même possible de transformer ce système en un modèle à observations complètes. Un résultat classique de la théorie du filtrage (voir Liptser & Shiriaev (1977)) le permet :

Proposition 5. *L'actif risqué S du modèle* (1.1)-(1.2) *est également solution de l'équation différentielle stochastique suivante :*

$$\frac{dS_t}{S_t} = E\left[\mu_t|\mathcal{F}_t^S\right]dt + \sigma_S dN_t, \tag{1.22}$$

où N est un $\left(\mathbb{P}, \mathbf{F}^S\right)$ mouvement Brownien.

Ce résultat donne une seconde justification théorique à l'utilisation du filtrage de Kalman faite dans le premier chapitre. D'ailleurs, si l'on considère le filtre de Kalman en temps continu et en régime permanent obtenu dans l'équation (1.6) du premier chapitre, il est possible d'obtenir une autre représentation de celui-ci :

Proposition 6. *En utilisant l'équation (1.6), on a :*

$$d\hat{\mu}_t = -\lambda\hat{\mu}_t dt + \lambda\sigma_S\left(\beta - 1\right)dN_t. \tag{1.23}$$

Le fait que le filtre de Kalman soit également un processus Gaussien est déjà connu. Ici, nous remarquons que, dans le régime permanent, le filtre de Kalman en temps continu est un processus d'Ornstein Uhlenbeck comme le drift. Dans la suite de ce chapitre, on va supposer que l'ensemble des paramètres du modèle (1.1)-(1.2) sont connus.

1.3.2 Performance de la stratégie optimale dans le cadre d'une tendance observable

Considérons le marché financier défini dans le premier chapitre avec un taux sans risque nul et sans coûts de transaction. Soit P^o le portefeuille autofinancé suivant :

$$\begin{aligned} \frac{dP_t^o}{P_t^o} &= \omega_t^o \frac{dS_t}{S_t}, \\ P_0^o &= x, \end{aligned}$$

où ω_t^o est la fraction de richesse investie dans l'actif risqué. Supposons également que l'agent souhaite maximiser sont utilité logarithmique espérée (avec une horizon T) dans un domaine admissible $\mathcal{A}^o$ pour l'allocation ω_t^o. Dans cette partie, nous supposons que l'agent est capable d'observer le processus μ. Formellement, cela revient à dire que $\mathcal{A}^o$ représente tous les processus adaptés et mesurables par rapport à la filtration $\mathbf{F}$. Dans ce cas, la solution de ce problème est donnée par (voir Lakner (1998) ou Bjork *et al.* (2010) par exemple) :

$$\begin{aligned} \frac{dP_t^o}{P_t^o} &= \frac{\mu_t}{\sigma_S^2} \frac{dS_t}{S_t}, && (1.24) \\ P_0^o &= x. && (1.25) \end{aligned}$$

Dans ce cas, il est possible d'obtenir l'équation différentielle stochastique des rendements logarithmiques :

Proposition 7. *Considérons le portefeuille P^o défini par l'équation (1.24). Dans ce cas,*

$$d\ln(P_t^o) = \frac{\mu_t^2}{2\sigma_S^2} dt + \frac{\mu_t}{\sigma_S} dW_t^S. \tag{1.26}$$

En utilisant cette équation, il est alors possible de caractériser la performance de cette stratégie :

Theorem 1.3.1. *Considérons le portefeuille P^o défini par l'équation (1.24). Dans ce cas,*

$$\begin{aligned} \lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T^o}{P_0^o}\right)\right]}{T} &= \frac{SNR}{2}, && (1.27) \\ \lim_{T\to\infty} \frac{\mathbb{V}ar\left[\ln\left(\frac{P_T^o}{P_0^o}\right)\right]}{T} &= SNR, && (1.28) \\ SR_\infty^o &= \frac{\sqrt{SNR}}{2}. && (1.29) \end{aligned}$$

où SNR est la ratio signal-sur-bruit :

$$SNR = \frac{\sigma_\mu^2}{2\lambda\sigma_S^2}. \tag{1.30}$$

Ce théorème nous montre que l'espérance et la variance asymptotique des rendements logarithmiques sont des fonctions linéaires du ratio signal-sur-bruit. Il nous montre également que le ratio de Sharpe est une fonction linéaire du ratio entre l'écart-types de la tendance et la volatilité.

1.3.3 Performance de la stratégie optimale dans le cadre d'une tendance non-observable

Considérons maintenant le même marché financier et un portefeuille autofinancé P tenu par un autre agent :

$$\begin{aligned} \frac{dP_t}{P_t} &= \omega_t \frac{dS_t}{S_t}, \\ P_0 &= x, \end{aligned}$$

où ω_t est aussi la fraction de richesse investie dans l'actif risqué. Cet agent cherche à résoudre le même problème d'optimisation mais n'est pas capable d'observer la tendance. Dans ce cas, l'agent recherche son allocation dans l'espace $\mathcal{A}$, qui représente cette fois tous les processus adaptés et mesurables par rapport à la filtration $\mathbf{F}^S$. La solution à ce problème est également connue (voir Lakner (1998) par exemple) :

$$\omega_t^* = \frac{E\left[\mu_t | \mathcal{F}_t^S\right]}{\sigma_S^2}.$$

En régime permanent, le portefeuille optimal devient alors :

$$\begin{aligned} \frac{dP_t}{P_t} &= \frac{\widehat{\mu}_t}{\sigma_S^2} \frac{dS_t}{S_t}, && (1.31) \\ P_0 &= x, && (1.32) \end{aligned}$$

où $\widehat{\mu}$ est défini dans l'équation (1.6). Il est également possible d'obtenir l'équation différentielle stochastique des rendements logarithmiques pour ce portefeuille :

Proposition 8. *Le portefeuille de l'équation (1.31) est solution de l'équation différentielle stochastique suivante :*

$$d\ln(P_t) = \frac{1}{2\sigma_S^2 \lambda(\beta - 1)} d\widehat{\mu}_t^2 + \left[\frac{\widehat{\mu}_t^2}{\sigma_S^2}\left(\frac{\beta}{(\beta-1)} - \frac{1}{2}\right) - \frac{1}{2}\lambda(\beta - 1)\right] dt,$$

où β est défini dans l'équation (1.7).

En utilisant cette équation, il est alors possible de caractériser la performance de cette stratégie :

Theorem 1.3.2. *Considérons le portefeuille de l'équation (1.31). Dans ce cas :*

$$\begin{aligned}
\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} &= \frac{\lambda}{4}\left(\beta-1\right)^2, && (1.33)\\
\lim_{T\to\infty} \frac{\mathbb{V}ar\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} &= \frac{\lambda}{8}\left(\beta^2-1\right)^2, && (1.34)\\
\lim_{T\to\infty} SR_T &= \sqrt{\frac{\lambda}{2}}\frac{\beta-1}{\beta+1}, && (1.35)
\end{aligned}$$

où β est défini dans l'équation (1.7).

De ce théorème, il est possible d'en extraire un corollaire qui caractérise les performances de ce portefeuille avec la vitesse de retour à la moyenne de la tendance et avec le ratio signal-sur-bruit :

Corollary 1.3.3. *Considérons le portefeuille de l'équation (1.31). Dans ce cas :*

$$\begin{aligned}
\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} &= \frac{1}{2}\left(SNR+\lambda-\sqrt{\lambda\left(\lambda+2SNR\right)}\right), && (1.36)\\
\lim_{T\to\infty} \frac{\mathbb{V}ar\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} &= \frac{SNR^2}{2\lambda}, && (1.37)\\
\lim_{T\to\infty} SR_T &= \left(\frac{\lambda}{2}\right)^{3/2}\frac{\left(\sqrt{1+\frac{2SNR}{\lambda}}-1\right)^2}{SNR}, && (1.38)
\end{aligned}$$

où SNR est la ratio signal-sur-bruit défini dans l'équation (1.30). De plus :

1. *Pour un paramètre λ fixé,*
   - *L'espérance asymptotique des rendements logarithmiques est une fonction croissante du ratio signal-sur-bruit,*
   - *le ratio de Sharpe asymptotique est également une fonction croissante du ratio signal-sur-bruit.*
2. *Pour un ratio signal-sur-bruit fixé,*
   - *L'espérance asymptotique des rendements logarithmiques est une fonction décroissante de λ,*
   - *le ratio de Sharpe asymptotique est une fonction décroissante de λ si :*

$$SNR < \frac{3}{2}\lambda, \tag{1.39}$$

et c'est une fonction croissante de λ si $SNR > \frac{3}{2}\lambda$.

Le ratio de Sharpe asymptotique maximum est atteint pour $\lambda = \frac{2}{3}SNR$ et vaut :

$$SR_{\infty}^{Max} = \frac{\sqrt{SNR}}{3^{3/2}}. \tag{1.40}$$

1.3.4 Impact de la non-observabilité de la tendance sur la stratégie optimale

Ici, nous cherchons a quantifier l'impact sur la performance de la non-observabilité de la tendance. Pour ce faire, nous introduisons le facteur d'observabilité partielle (partial information factor en anglais). Cet indicateur est défini comme étant le ratio de Sharpe de la stratégie optimale avec informations partielles (tendance non observable) divisé par le ratio de Sharpe de la stratégie optimale avec informations complètes (tendance observée) :

$$\mathrm{PIF} = \frac{\mathrm{SR}_{\infty}}{\mathrm{SR}_{\infty}^{\mathrm{o}}}, \tag{1.41}$$

En utilisant les parties précédentes, il est alors possible d'énoncer le théorème suivant :

Theorem 1.3.4. *Le facteur d'observabilité partielle est égal à :*

$$PIF = \left(\frac{\lambda}{SNR}\right)^{3/2} \frac{\left(\sqrt{1+\frac{2SNR}{\lambda}}-1\right)^2}{\sqrt{2}}, \tag{1.42}$$

où SNR est la ratio signal-sur-bruit défini dans l'équation (1.30). Si $SNR < \frac{3}{2}\lambda$ (respectivement, $SNR > \frac{3}{2}\lambda$):

1. *Pour un ratio signal-sur-bruit fixé, cet indicateur est une fonction décroissante (respectivement croissante) de λ.*

2. *Pour un paramètre λ fixé, cet indicateur est une fonction croissante (respectivement décroissante) du ratio signal-sur-bruit.*

De plus :

$$PIF \leq \frac{2}{3^{3/2}}, \tag{1.43}$$

et cette borne est atteinte pour $\lambda = \frac{2}{3}SNR$.

Ce théorème nous montre qu'au mieux, le ratio de Sharpe asymptotique de la stratégie optimal avec informations partielles est approximativement égal à 38.49% du ratio de Sharpe de la stratégie optimale avec informations complètes.

De plus, l'intuition nous dirait qu'un ratio signal-sur-bruit élevé et une faible vitesse de retour à la moyenne de la tendance λ implique un impact faible de l'observabilité de la tendance sur la performance de la stratégie optimale (et donc un facteur d'observabilité partielle élevé). Cette intuition ne se vérifie que si $\mathrm{SNR} \leq \frac{3}{2}\lambda$.

1.3.5 Conclusion

Ce chapitre quantifie l'impact de la non-observabilité de la tendance sur la performance de la stratégie optimale avec un modèle où le drift est un processus caché d'Ornstein Uhlenbeck.
Si la tendance est observable, nous avons montré que le ratio de Sharpe asymptotique est uniquement une fonction croissante du ratio signal-sur-bruit.
Sous observations partielles, ce ratio de Sharpe asymptotique devient une fonction du ratio signal-sur-bruit et de la vitesse de retour à la moyenne de la tendance. Même s'il demeure une fonction croissante du ratio signal-sur-bruit, nous avons trouvé que ce n'est pas une fonction monotone de la vitesse de retour à la moyenne de la tendance (croissante puis décroissante).
Nous avons également montré que le ratio de Sharpe asymptotique de la stratégie optimale avec informations partielles ne peut pas dépasser $\frac{2}{3^{3/2}} *$ 100% du ratio de Sharpe asymptotique de la stratégie optimale avec informations complètes.
Nous avons également montré que malgré un ratio signal-sur-bruit élevé, un retour à la moyenne trop rapide de la tendance implique une performance négligeable de la stratégie avec observations partielles comparée à celle avec observations complètes.

1.4 Robustesse de la stratégie optimale avec des paramètres mal-spécifiés et d'une stratégie d'analyse technique

Nous pouvons distinguer plusieurs types d'investissement dans l'industrie financière (voir Blanchet-Scalliet *et al.* (2007)) tel que l'analyse fondamentale (voir Tideman (1972)), l'analyse technique (voir Taylor & Allen (1992), Brown & Jennings (1989) et Edwards *et al.* (2007) pour plus de détails) et l'approche mathématique introduite par Merton en 1969 (voir Merton (1990) pour plus de détails).
Comme nous l'avons vu dans le premier chapitre, l'approche de type modèle est confrontée au problème de calibration sur les données réelles. En effet, le faible ratio signal-sur-bruit présent sur les séries financières ne permet pas d'avoir une bonne calibration de la tendance. Le second chapitre nous a permis d'illustrer l'impact de la non-observabilité du drift. Nous allons donc considérer dans ce chapitre la stratégie optimale avec informations partielles introduite dans le second chapitre en supposant que l'investisseur se trompe dans la calibration de sa tendance. Le but est d'étudier l'impact d'une erreur de calibration sur la performance de ce type d'investissement.
Afin d'évaluer la pertinence de cette stratégie, nous allons également étudier un investissement basé sur de l'analyse technique pour pouvoir comparer

les performances et la robustesse des deux stratégies. Blanchet-Scalliet *et al.* (2007) se sont également intéressés à ce genre de problème. En effet, leur étude considère le cas d'un drift continu par morceaux, non-observable et qui change de valeur à un instant aléatoire. Après avoir caractérisé l'investissement optimal dans le cadre d'une mauvaise calibration, ils ont ensuite utilisé des simulations de type Monte Carlo pour montrer qu'une stratégie utilisant une moyenne arithmétique glissante peut avoir une meilleure performance que la stratégie optimale mal calibrée. Avec le modèle introduit dans le premier chapitre, Zhu & Zhou (2009) ont considéré une stratégie basée sur une comparaison entre la valeur de l'actif risqué et sa moyenne géométrique glissante. Dans ce chapitre, la stratégie que nous considérons est basée sur la comparaison entre deux moyennes géométriques glissantes : une court terme et une long terme.
Pour effectuer cette analyse, nous allons donc dans un premier temps caractériser, de façon analytique, la performance (caractérisée par l'espérance asymptotique du rendement logarithmique) de la stratégie optimale mal-spécifiée et celle d'un investissement utilisant des croisements de moyennes mobiles géométriques. Après avoir obtenu ces résultats, nous montrerons que sur plusieurs types de régimes la stratégie utilisant l'analyse technique s'avère être plus robuste que la stratégie optimale mal calibrée, et cela même avec une volatilité stochastique.

1.4.1 Préliminaires

Dans ce chapitre, nous considérons toujours le modèle (1.1)-(1.2) introduit dans le premier chapitre. A titre de rappel, il a été montré dans celui-ci que le filtre de Kalman en temps continu et en régime permanent est solution de l'équation différentielle stochastique suivante :

$$d\widehat{\mu}_t = -\lambda\beta\widehat{\mu}_t dt + \lambda\left(\beta - 1\right)\frac{dS_t}{S_t},$$

où

$$\beta = \left(1 + \frac{\sigma_\mu^2}{\lambda^2\sigma_S^2}\right)^{\frac{1}{2}}.$$

Nous pouvons montrer également que ce filtre peut être réécrit comme une moyenne mobile exponentielle des rendements corrigée par un facteur :

Proposition 9.

$$\widehat{\mu}_t = m^*\tilde{\mu}_t^*, \tag{1.44}$$

où $m^* = \frac{\beta-1}{\beta}$ *et* $\tilde{\mu}^*$ *est une moyenne mobile exponentielle vérifiant :*

$$d\tilde{\mu}_t^* = -\frac{1}{\tau^*}\tilde{\mu}_t^* dt + \frac{1}{\tau^*}\frac{dS_t}{S_t}, \tag{1.45}$$

avec un temps caractéristique valant $\tau^* = \frac{1}{\lambda\beta}$.

1.4.2 Stratégie optimale avec des paramètres mal-spécifiés

Considérons le marché financier introduit dans le second chapitre. Soit P un portefeuille autofinancé vérifiant :

$$\begin{aligned} \frac{dP_t}{P_t} &= \omega_t \frac{dS_t}{S_t}, \\ P_0 &= x, \end{aligned}$$

où ω_t représente le pourcentage de richesse investi dans l'actif risqué. Comme nous l'avions vu dans le second chapitre, si l'agent ne peux pas observer la tendance et qu'il souhaite maximiser son utilité logarithmique espérée (avec une horizon T), il doit utiliser l'allocation suivante :

$$\omega_t^* = \frac{E\left[\mu_t | \mathcal{F}_t^S\right]}{\sigma_S^2}.$$

Plaçons nous maintenant en régime permanent et supposons que l'agent effectue une mauvaise calibration du modèle (1.1)-(1.2). Vu la proposition 9, le fait de se tromper sur les valeurs (λ, σ_μ) est équivalent au fait de se tromper sur le facteur $\frac{\beta-1}{\beta}$ et le temps caractéristique τ^*. C'est pour cette raison que nous considérons que l'agent pense que le temps caractéristique vaut τ :

$$d\tilde{\mu}_t = -\frac{1}{\tau}\tilde{\mu}_t dt + \frac{1}{\tau}\frac{dS_t}{S_t}, \tag{1.46}$$

$$\tilde{\mu}_0 = 0, \tag{1.47}$$

et utilise l'allocation suivante :

$$\frac{dP_t}{P_t} = m\frac{\tilde{\mu}_t}{\sigma_S^2}\frac{dS_t}{S_t}, \tag{1.48}$$

$$P_0 = x, \tag{1.49}$$

où $m > 0$. Il est alors possible de formuler la propriété suivante :

Proposition 10. *Le portefeuille P de l'équation (1.48) est solution de :*

$$d\ln(P_t) = \frac{m\tau}{2\sigma_S^2} d\tilde{\mu}_t^2 + m\left(\frac{\tilde{\mu}_t^2}{\sigma_S^2}\left(1 - \frac{m}{2}\right) - \frac{1}{2\tau}\right) dt. \tag{1.50}$$

Partant de ce résultat, il est possible d'énoncer le théorème suivant :

Theorem 1.4.1. *Considérons le portefeuille P de l'équation (1.48). Dans ce cas :*

$$\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = m\frac{\tau\left(\beta^2-1\right)(2-m) - m\left(\tau+\frac{1}{\lambda}\right)}{4\tau\left(\tau+\frac{1}{\lambda}\right)}, \tag{1.51}$$

où β est défini dans l'équation (1.7).

De ce théorème, il est possible d'en extraire un corollaire qui caractérise la performance de ce portefeuille avec la vitesse de retour à la moyenne de la tendance et avec le ratio signal-sur-bruit :

Corollary 1.4.2. *Considérons le portefeuille P de l'équation (1.48). Dans ce cas :*

$$\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = m\frac{2\tau(2-m)\,SNR - m(\lambda\tau+1)}{4\tau(\lambda\tau+1)}, \tag{1.52}$$

où SNR est la ratio signal-sur-bruit défini dans l'équation (1.30). De plus:

1. *Si $m < 2$, pour un paramètre λ fixé, L'espérance asymptotique des rendements logarithmiques est une fonction croissante du ratio signal-sur-bruit.*

2. *Pour un ratio signal-sur-bruit fixé, c'est une fonction décroissante de λ.*

En utilisant ce résultat, nous pouvons énoncer la proposition suivante qui nous informe quelles contraintes doivent être respectées pour que la performance de la stratégie optimale mal spécifiée soit positive et qu'il existe un temps caractéristique optimal (maximisant la performance) :

Proposition 11. *Considérons le portefeuille P de l'équation (1.48) avec $m < 2$. Dans ce cas, l'espérance asymptotique du rendement logarithmique est positive si, et seulement si :*

1. $\frac{SNR}{\lambda} > \frac{2m}{2-m}$.

2. $\tau > \tau_{min}$, *où:*

$$\tau_{min} = \frac{m}{2(2-m)\,SNR - \lambda m}. \tag{1.53}$$

De plus, il existe un temps caractéristique optimal $\tau_{min} < \tau_{opt} < \infty$ si, et seulement si $\frac{SNR}{\lambda} > \frac{2m}{2-m}$ et :

$$\tau_{opt} = \frac{m + \sqrt{(2-m)\,2m\frac{SNR}{\lambda}}}{2(2-m)\,SNR - \lambda m}. \tag{1.54}$$

1.4.3 Investissement utilisant des croisements de moyennes mobiles géométriques

Nous considérons le même marché financier introduit dans le chapitre deux. Soit $G(t, L)$ la moyenne géométrique à l'instant t sur une fenêtre L des prix de l'actif risqué :

$$G(t, L) = \exp\left(\frac{1}{L}\int_{t-L}^{t} \log(S_u)\, du\right), \tag{1.55}$$

et soit Q un portefeuille autofinancé vérifiant :

$$\frac{dQ_t}{Q_t} = \theta_t \frac{dS_t}{S_t}, \tag{1.56}$$

$$Q_0 = x, \tag{1.57}$$

où θ_t représente le pourcentage de richesse investi dans l'actif risqué :

$$\theta_t = \gamma + \alpha\, \mathbf{1}_{G(t,L_1)>G(t,L_2)}$$

avec $\gamma, \alpha \in \mathbb{R}$ et $0 < L_1 < L_2 < t$. On suppose qu'avant l'instant L_2, l'allocation est nulle. l'introduction des deux coefficients γ, α permet de considérer une stratégie de type "suiveur de tendance" ou de type "retour à la moyenne". En effet, on peut choisir soit d'acheter, soit de vendre lorsque la moyenne court terme est plus grande que la moyenne long terme. Il est alors possible de formuler la proposition suivante :

Proposition 12. *Le portefeuille P de l'équation (1.56) est solution de :*

$$\begin{aligned} d\ln(Q_t) \quad = \quad & \left(\left(\gamma + \alpha\, \mathbf{1}_{G(t,L_1)>G(t,L_2)}\right)\mu_t - \frac{\gamma^2\sigma_S^2}{2}\right. \\ & \left. - \frac{(\alpha^2 + 2\alpha\gamma)\,\sigma_S^2}{2}\, \mathbf{1}_{G(t,L_1)>G(t,L_2)}\right) dt \\ & + \left(\gamma + \alpha\, \mathbf{1}_{G(t,L_1)>G(t,L_2)}\right)\sigma_S dW_t^S. \end{aligned}$$

A l'aide de cette proposition, nous pouvons alors formuler le théorème suivant qui fournit la performance asymptotique de cette stratégie :

Theorem 1.4.3. *Considérons le portefeuille P de l'équation (1.56). Dans ce cas:*

$$\begin{aligned} \lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{Q_T}{Q_0}\right)\right]}{T} = & -\frac{\gamma^2\sigma_S^2}{2} - \frac{(\alpha^2 + 2\alpha\gamma)\,\sigma_S^2}{2}\Phi\left(\frac{m_{(L_1,L_2,\sigma_S)}}{\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\right) \\ + & \frac{\alpha\sigma_\mu^2\left(L_2\left(1-e^{-\lambda L_1}\right) - L_1\left(1-e^{-\lambda L_2}\right)\right)}{2\lambda^3 L_1 L_2 \sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\Phi'\left(-\frac{m_{(L_1,L_2,\sigma_S)}}{\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\right), \end{aligned}$$

où Φ représente la fonction de répartition de la loi normale centrée réduite et :

$$
\begin{aligned}
m_{(L_1,L_2,\sigma_S)} &= \frac{-\sigma_S^2}{4}(L_2 - L_1), \\
s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)} &= \left(\frac{\sigma_\mu^2}{\lambda^2} + \sigma_S^2\right)\frac{(L_2 - L_1)^2}{3L_2} - \frac{\sigma_\mu^2}{\lambda^4}\left(\frac{1}{L_1} - \frac{1}{L_2}\right) \\
&\quad + \frac{\sigma_\mu^2}{\lambda^5}\left[\frac{1}{L_1^2}\left(1 - e^{-\lambda L_1}\right) + \frac{1}{L_2^2}\left(1 - e^{-\lambda L_2}\right)\right. \\
&\quad - \frac{1}{L_1 L_2}\left(1 - e^{-\lambda L_1}\right)\left(1 - e^{-\lambda L_2}\right) \\
&\quad \left. - \frac{1}{L_1 L_2}\left(e^{-\lambda(L_2 - L_1)} - e^{-\lambda(L_2 + L_1)}\right)\right].
\end{aligned}
$$

1.4.4 Robustesse de la stratégie optimale avec des paramètres mal calibrés et d'une stratégie d'analyse technique

Cas d'une volatilité constante

Après avoir caractérisé la performance asymptotique de chacune de ces deux stratégies, nous pouvons alors effectuer une étude de robustesse en considérant deux exemples d'applications de ces allocations. Pour ce faire, nous fixons le nombre de jours de trading par an à 252. Nous fixons également la volatilité de l'actif risqué à $\sigma_S = 30\%$. Nous considérons alors la stratégie optimale mal-spécifiée avec un levier $m = 1$ et un temps caractéristique $\tau = 252$ jours. La stratégie utilisant le croisement de moyennes mobiles géométriques est quant à elle prise avec $(L_1, L_2) = (5 \text{ jours}, 252 \text{ jours})$. Cela revient à comparer la moyenne d'une semaine avec celle prise sur une année. Dans cet exemple, nous supposons que si la moyenne court terme est supérieure (respectivement inférieure) à la moyenne long terme, l'agent achète (respectivement vend) l'actif risqué. Formellement, l'allocation est donnée par :

$$\theta_t = -1 + 2\,\mathbf{1}_{G(t,L_1) > G(t,L_2)}.$$

Pour comparer ces stratégies, nous utilisons les résultats des théorèmes 1.4.1 et 1.4.3. Les figures suivantes représentent les espérances asymptotiques des rendements logarithmiques obtenues après 100 années d'investissement en fonction de la volatilité de la tendance σ_μ avec respectivement $\lambda = 1, 2, 3$ et 4. Même si la stratégie optimal peut fournir une meilleure performance dans certaines configurations (par exemple pour $\lambda = 1$ et $\sigma_\mu = 90\%$), elle peut aussi subir une plus grosse perte (par exemple pour $\lambda = 4$ et $\sigma_\mu = 10\%$). Sur toutes les configurations étudiées, la stratégie d'analyse technique présente une performance plus stable que la stratégie optimale, c'est en ce sens qu'elle est plus robuste.

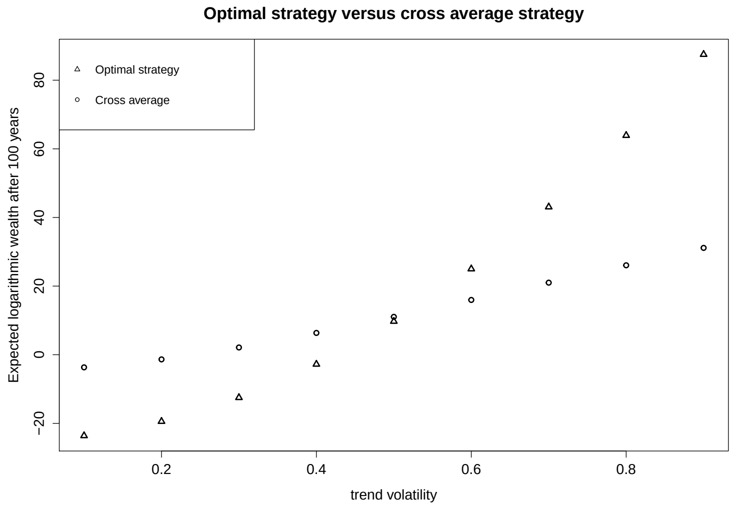

Figure 1.1: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 1$, $\sigma_S = 30\%$ and $T = 100$ years

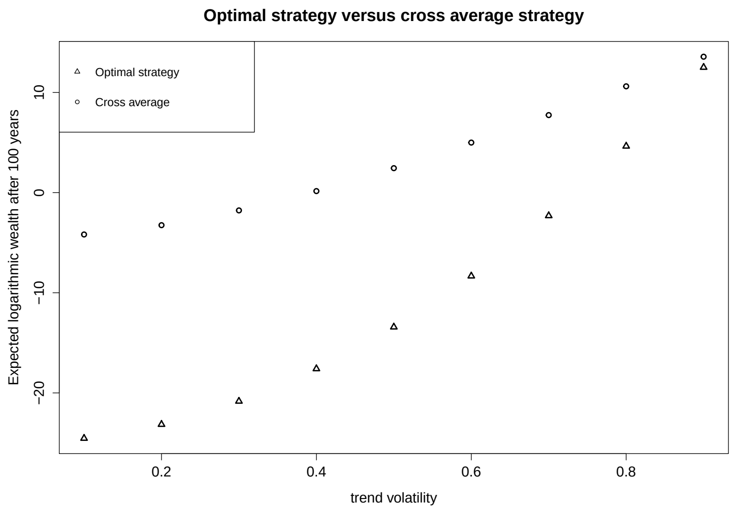

Figure 1.2: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 2$, $\sigma_S = 30\%$ and $T = 100$ years

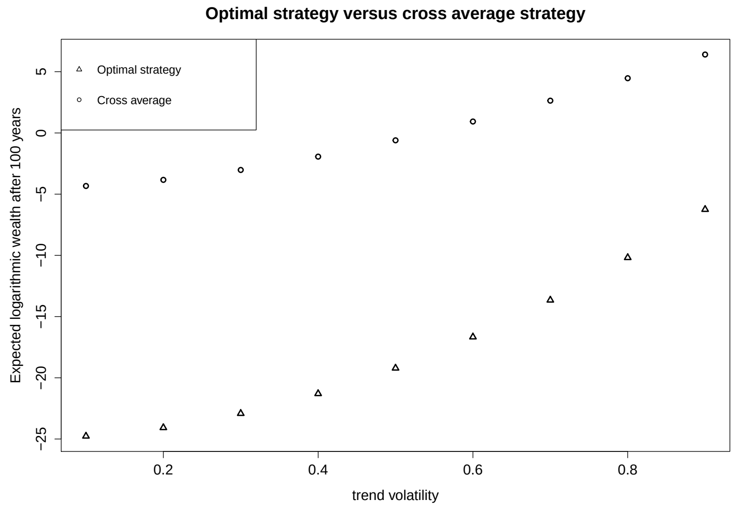

Figure 1.3: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 3$, $\sigma_S = 30\%$ and $T = 100$ years

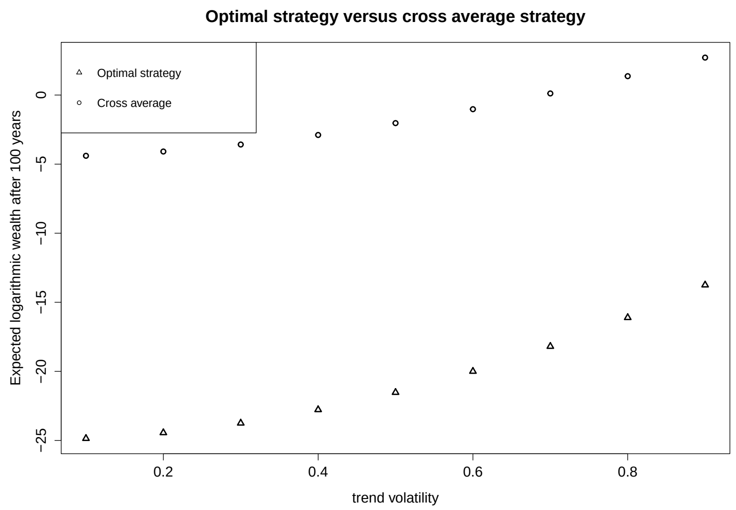

Figure 1.4: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 4$, $\sigma_S = 30\%$ and $T = 100$ years

Cas d'une volatilité stochastique de Heston

Après avoir utilisé les résultats analytiques pour étudier la robustesse de ces deux stratégies dans le cas d'une volatilité constante, cette sous-partie s'intéresse au cas d'une volatilité stochastique de Heston (voir Heston (1993) ou Mikhailov & Nögel (2003) pour plus de détails) et répond à cette question grâce à des simulations de type Monte Carlo. Pour ce faire, considérons un marché financier défini sur un espace de probabilité $(\Omega, \mathcal{G}, \mathbf{G}, \mathbb{P})$, où $\mathbf{G} = \{\mathcal{G}_t, t \geqslant 0\}$ est la filtration naturelle engendrée par l'association de trois processus Browniens (W^S, W^μ, W^V), et $\mathbb{P}$ la mesure de probabilité historique. Le modèle d'actif risqué suivant est alors supposé :

$$\begin{aligned}
\frac{dS_t}{S_t} &= \mu_t dt + \sqrt{V_t} dW_t^S, \\
d\mu_t &= -\lambda \mu_t dt + \sigma_\mu dW_t^\mu, \\
dV_t &= \alpha \left(V_\infty - V_t\right) dt + \epsilon \sqrt{V_t} dW_t^V
\end{aligned}$$

avec $\mu_0 = 0$, $V_0 > 0$, $d\left\langle W^S, W^\mu \right\rangle_t = 0$, et $d\left\langle W^S, W^V \right\rangle_t = \rho dt$. Nous supposons également que $(\lambda, \sigma_\mu) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ et que $2kV_\infty > \epsilon$ (cette condition assure le fait que la volatilité soit toujours strictement positive, voir Cox *et al.* (1985) pour plus de détails). Notons également $\mathbf{G}^S = \left\{\mathcal{G}_t^S\right\}$ la filtration engendrée par S. Dans ce cas, le processus V est adapté à la filtration G^S (en effet, l'utilisation du crochet sur $\ln S$ peut le montrer). Dans ce cas, si l'investisseur cherche à résoudre le même problème d'investissement optimal que précédemment (maximisation de sa richesse logarithmique espérée future avec une allocation G^S-mesurable et avec un portefeuille autofinancé), il investira de cette façon (voir Bjork *et al.* (2010) pour plus d'informations) :

$$\begin{aligned}
\frac{dP_t}{P_t} &= \frac{E\left[\mu_t | \mathcal{G}_t^S\right]}{V_t} \frac{dS_t}{S_t}, \\
P_0 &= x.
\end{aligned}$$

Pour effectuer nos simulations, nous nous sommes placés en temps discret. Si l'on note δ la fréquence d'échantillonnage de notre système, et que l'on représente l'instant $t_k = k\delta$ par l'indice k, la version en temps discret de ce modèle qui introduit le plus petit biais sur le processus de variance V (voir Lord *et al.* (2010) pour une étude sur la discrétisation du modèle de Heston) peut alors s'écrire :

$$y_{k+1} = \frac{S_{k+1} - S_k}{\delta S_k} = \mu_{k+1} + u_{k+1}, \tag{1.58}$$

$$\mu_{k+1} = e^{-\lambda\delta} \mu_k + v_k, \tag{1.59}$$

$$V_{k+1} = V_k + \alpha \left(V_\infty - V_k^+\right) \delta + \epsilon \sqrt{V_k^+} z_k \tag{1.60}$$

où $x^+ = \max(0, x)$, $u_{k+1} \sim \mathcal{N}\left(0, \frac{V_k}{\delta}\right)$, $v_k \sim \mathcal{N}\left(0, \frac{\sigma_\mu^2}{2\lambda}\left(1 - e^{-2\lambda\delta}\right)\right)$ et $z_k \sim \mathcal{N}(0, \delta)$. Nous avons alors supposé pour nos simulations que $\epsilon = 5\%$, que $V_\infty = V_0 = 0.3^2$ (ce qui revient à dire que la volatilité initiale et la volatilité "long terme" valent 30%) et que $\rho = -60\%$ (lorsque le prix de l'actif risqué baisse, la volatilité augmente). De plus, nous nous sommes fixés une horizon de 50 ans avec un pas d'échantillonnage journalier $\delta = 1/252$ (en admettant qu'il y a 252 jours dans une année) et une ré-allocation journalière des deux stratégies suivantes:

1. La stratégie optimale discrétisée. Vu que le processus V est adapté à la filtration G^S, V_k est connu à l'instant t_k, et dans ce cas le fitre de Kalman non stationnaire en temps discret peut être implémenté. Nous supposons également que l'agent croit que les paramètres de la tendance valent $\lambda^a = 1$ et $\sigma_\mu^a = 90\%$.

2. La stratégie utilisant le croisement de moyennes mobiles géométriques est quant à elle prise avec $(L_1, L_2) = (5 \text{ jours}, 252 \text{ jours})$. Nous supposons que si la moyenne court terme est supérieure (respectivement inférieure) à la moyenne long terme, l'agent achète (respectivement vend) l'actif risqué.

Les deux figures suivantes représentent les espérances empiriques des rendements logarithmiques obtenues après 50 années d'investissement et avec 10000 trajectoires en fonction de la volatilité de la tendance σ_μ avec respectivement $\lambda = 1$ et 2. Nous obtenons des résultats similaires à ceux obtenus dans le cas d'une volatilité constante. Même si la stratégie optimale bien calibrée à une meilleur espérance des rendements logarithmiques, la stratégie d'analyse technique a une performance plus stable lorsque la dynamique de la tendance varie. Les quatre dernières figures représentent les distributions empiriques du rendement logarithmique après 50 années et sur 10000 trajectoires pour différentes dynamiques de tendance. Nous remarquons que même dans le cas bien calibré, la stratégie optimale est plus dispersée que la stratégie d'analyse technique. Ce résultat n'est pas en contradiction avec la définition que l'on a donné à la stratégie optimale. Dans notre étude, nous avons considéré la stratégie maximisant la richesse logarithmique espérée, le résultat ne minimise donc en aucune façon la dispersion de la distribution de cette richesse finale. Il est donc plus robuste d'investir en utilisant des croisements de moyennes mobiles.

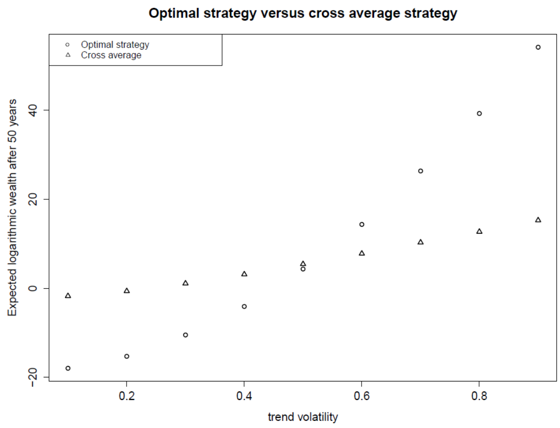

Figure 1.5: The expected logarithmic returns of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $M = 10000$, $\lambda = 1$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

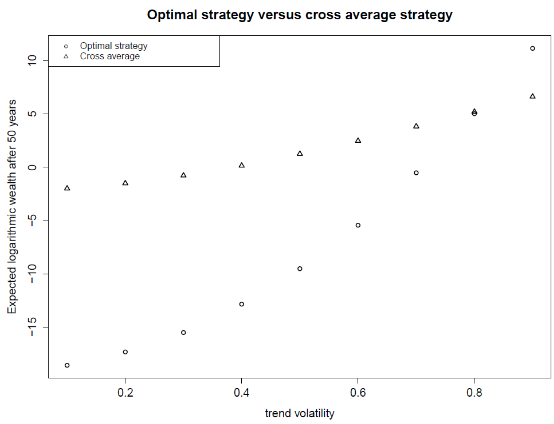

Figure 1.6: The expected logarithmic returns of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days)as functions of σ_μ with $M = 10000$, $\lambda = 2$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

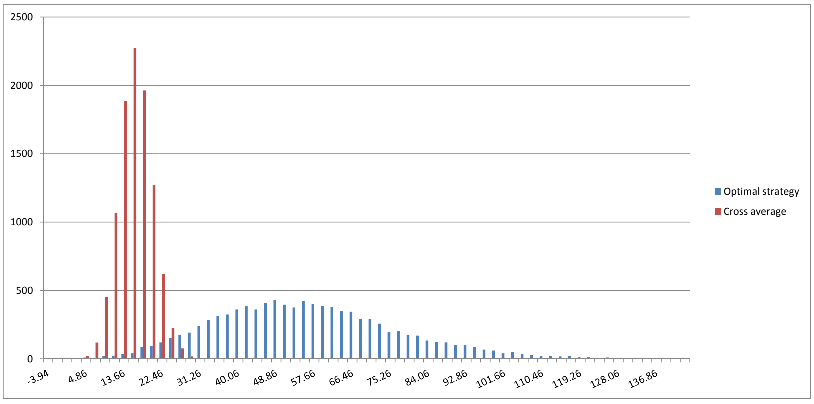

Figure 1.7: Empirical distribution of the logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 90\%$, $\lambda = 1$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

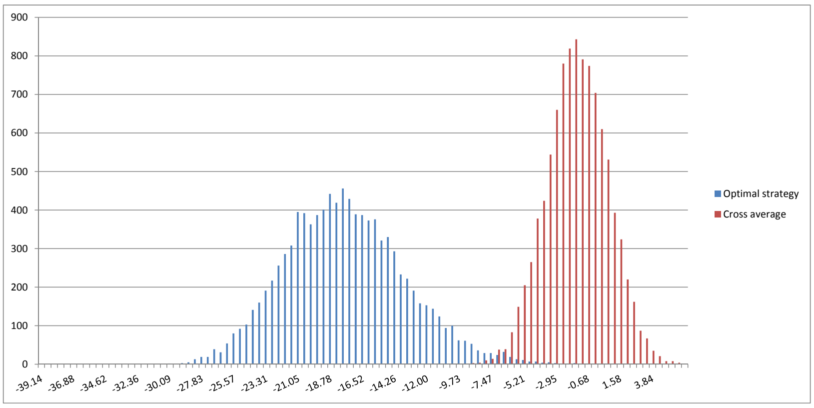

Figure 1.8: Empirical distribution of the expected logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 10\%$, $\lambda = 1$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

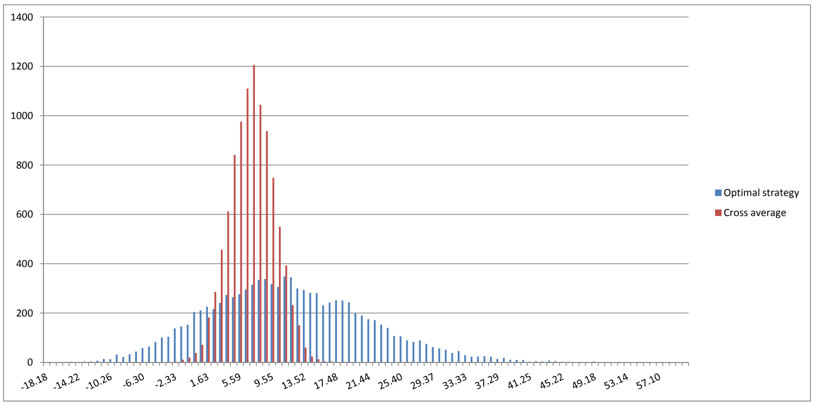

Figure 1.9: Empirical distribution of the expected logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 90\%$, $\lambda = 2$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

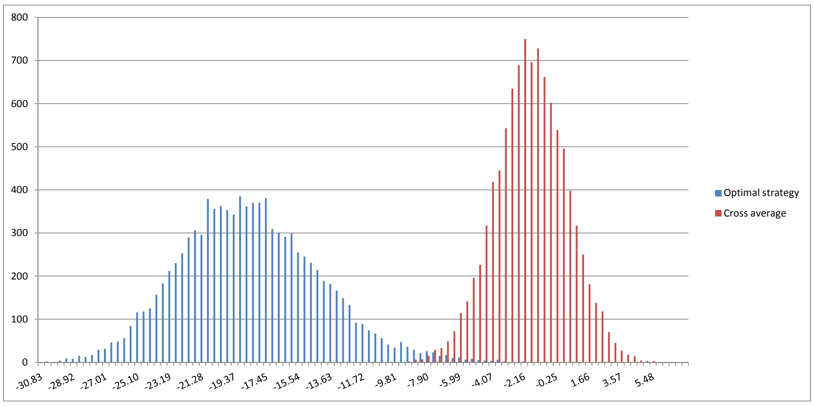

Figure 1.10: Empirical distribution of the expected logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 10\%$, $\lambda = 2$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

1.4.5 Conclusion

Ce chapitre a quantifié de façon analytique la performance de la stratégie optimale mal calibrée et celle d'une stratégie d'analyse technique avec un modèle où le drift est un processus caché d'Ornstein Uhlenbeck.
Pour la stratégie optimale, nous avons montré que l'espérance asymptotique du rendement logarithmique est une fonction croissante du ratio signal-sur-bruit et une fonction décroissante de la vitesse de retour à la moyenne de la tendance.
Nous avons également montré que, dans le cadre d'une mauvaise calibration, la performance asymptotique de cette stratégie pouvait être positive sous certaines conditions sur les paramètres du modèle et de la stratégie. C'est d'ailleurs sous ces mêmes conditions que nous avons montré qu'il existait un filtrage optimal qui coïncide avec le filtre de Kalman dans le cas d'une bonne calibration.
Concernant la stratégie basée sur des croisements de moyennes mobiles géométriques, nous avons également pu caractériser, de façon analytique, l'espérance asymptotique du rendement logarithmique.
De plus, les simulations que nous avons effectué montre que, sur un modèle où le drift est un processus caché d'Ornstein Uhlenbeck, et même avec une volatilité stochastique, l'analyse technique est plus robuste que la stratégie optimale mal calibrée.

Note

Chaque chapitre de cette thèse a été publié séparément. Nous avons gardé, volontairement, les versions originales des papiers. Ceci permet à chaque lecteur de comprendre parfaitement la partie qui l'intéresse sans besoin de lire les parties précédentes.

Each chapter of this thesis was published separately. We kept, deliberately, the original versions of the papers. This allows each reader to fully understand the part that interests him without need to read the previous parts

Chapter 2

Forecasting trends with asset prices

The question addressed in this chapter is the estimation of the trend of a financial asset, and the impact of misspecification. The setting we consider is that of a stochastic asset price model where the trend follows an unobservable Ornstein-Uhlenbeck process. Motivated by the use of Kalman filtering as a forecasting tool, we study the problem of parameters estimation, and measure the effect of parameters misspecification. Numerical examples illustrate the difficulty of trend forecasting in financial time series.

Introduction

Asset prices may be well described by random walks, as the Efficient Market Hypothesis advocates. If this is indeed the case, then future returns are not predictable. Nevertheless, professionals in the finance industry tend to have divergent views on the subject, and trend following strategies are the principal sources of returns for Commodity Trading Advisors (see Lempérière *et al.* (2014)). In fact, most quantitative strategies are based on the more or less explicit assumption that the trends of assets are known (see Markovitz (1952), Merton (1990)) and can be extracted from the asset prices themselves. It is therefore natural to address the question of forecasting asset trends, and to seek to provide reliable statistical estimators.
Unfortunately, the estimation of the trend of an asset is a statistically difficult problem, mainly because of a high measurement noise: consider for example a simple model with a constant trend $\frac{dS_t}{S_t} = \mu dt + \sigma_S dW_t^S$. Then, the best estimate of the trend at time T is given by $\widehat{\mu}_t = \frac{1}{T}\int_0^T \frac{dS_u}{S_u}$. Student's t-test will reject the hypothesis $\mu = 0$ if $|\widehat{\mu}_T| > \frac{1.96\sigma_S}{\sqrt{T}}$ at a 5% significance level. Therefore, with $\sigma_S = 30\%$, the estimate $\widehat{\mu}_T = 1\%$ becomes statistically relevant for observation times $T > 3457$ years...

The purpose of this work is to assess the feasibility of forecasting trends modelled by an unobserved mean-reverting diffusion.
Using a Bayesian approach or maximum likelihood estimation (see e.g. Leroux (1992),O. Cappé & Moulines (2005), Benmiloud & Pieczynski (1995), Casarin & Marin (2007) or Dahia (2005)), inference methods for unobservable processes have been applied to financial time series, mostly in the framework of stochastic volatility models (see Jacquier *et al.* (1994),Eraker (1998),Kim *et al.* (1998) or Chib *et al.* (2002)). Closer in spirit to our motivation, several authors have considered the situation of an unobservable stochastic trend, and use filtering methods (see Lakner (1998), Pham & Quenez (2001),Laskry & Lions (1999) or Brendle (2006a)) in this context. As it turns out, most of these filters are based on a parametric stochastic model for the trend, and their usefulness in realistic trading strategies is therefore confronted to the problem of parameters estimation. It is our aim to partly fill the existing gap in the quantitative finance literature, and shed a new light on the feasability of classical trading strategies based on the determination of asset trends.
The chapter is organized as follows: in Section 1, we present the model and recall some results from Kalman filtering. Section 2 is devoted to the inference of the parameters using discrete time observations. The performance of statistical estimators is evaluated by giving their asymptotic behaviours, and by providing, in closed form, the Cramer-Rao bound. Section 3 introduces the continuous time misspecified Kalman filter. We provide estimates for the impact of parameters misspecification on trend filtering, and compute the probability to have a positive trend, knowing a positive estimate. Finally, Section 4 contains numerical examples illustrating the relevance of parameters misspecification in trend filtering.

2.1 Framework

In this section, the model for the asset price and the mean-reverting dynamics of its trend is made precise. Then the Kalman filtering method, based on a time-discretized version, is recalled.

2.1.1 Model

Continuous time model

Consider a financial market living on a stochastic basis $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$, where $\mathbf{F} = \{\mathcal{F}_t, t \geqslant 0\}$ is the natural filtration associated to a two-dimensional (uncorrelated) Wiener process (W^S, W^μ), and $\mathbb{P}$ is the objective probability measure. The dynamics of the risky asset S is given by

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu_t dt + \sigma_S dW_t^S, && (2.1) \\ d\mu_t &= -\lambda \mu_t dt + \sigma_\mu dW_t^\mu, && (2.2) \end{aligned}$$

with $\mu_0 = 0$. We also assume that $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$.
Denote by $\mathbf{F}^S = \left\{ \mathcal{F}_t^S \right\}$ be the natural filtration associated to the price process S. An important point is that only $\mathbf{F}^S$-adapted processes are observable, which implies that agents in this market do not observe the trend μ.

Discrete time model

A discrete-time version of (2.1)-(2.2) is now presented. Let then δ be a discrete time step, and denote by the subscript k the value of a process at time $t_k = k\delta$. The discrete time model is:

$$\begin{aligned} y_{k+1} = \frac{S_{k+1} - S_k}{\delta S_k} &= \mu_{k+1} + u_{k+1}, && (2.3) \\ \mu_{k+1} &= e^{-\lambda\delta} \mu_k + v_k, && (2.4) \end{aligned}$$

where $u_k \sim \mathcal{N}\left(0, \frac{\sigma_S^2}{\delta}\right)$ and $v_k \sim \mathcal{N}\left(0, \frac{\sigma_\mu^2}{2\lambda}\left(1 - e^{-2\lambda\delta}\right)\right)$. The system (2.3)-(2.4) corresponds to an AR(1) model with noise.

2.1.2 Optimal trend estimator

Discrete Kalman filter

In this subsection, the parameters $\theta = (\lambda, \sigma_\mu)$ and σ_S are supposed to be known. The discrete time system (2.3)-(2.4) corresponds to a Linear Gaussian Space State model where the observation is y and the state, μ (see Brockwell & Davis (2002a) for details). In this case, the optimal estimator, given by the Kalman filter[1], is the conditional expectation $\mathbb{E}[\mu_k | y_1, ..., y_k]$. For simplicity, we let $\widehat{X}_{k/l}$ denote $\mathbb{E}[X_k | y_1, ..., y_l]$. The Kalman filter is decomposed in two distinct phases:

1. An *a priori* estimate given $\hat{\mu}_{k+1/k}$ and $\Gamma_{k+1/k} = \mathbb{E}\left[(\mu_{k+1} - \hat{\mu}_{k+1/k})(\mu_{k+1} - \hat{\mu}_{k+1/k})^T\right]$. This estimate is done using the transition equation (2.4).

2. An *a posteriori* estimate. When the new observation is available, a correction of the first estimate is done to obtain $\hat{\mu}_{k+1/k+1}$ and $\Gamma_{k+1/k+1} = \mathbb{E}\left[(\mu_{k+1} - \hat{\mu}_{k+1/k+1})(\mu_{k+1} - \hat{\mu}_{k+1/k+1})^T\right]$. The criterion for this correction is the least squares method.

[1] Appendix A presents a detailed introduction to the discrete Kalman filter

Thus, $\hat{\mu}_{k/k}$ is the minimum variance linear unbiased estimate of the trend μ_k. Formally, the iterative method is given by:

$$\widehat{\mu}_{k+1/k+1} = e^{-\lambda\delta}\widehat{\mu}_{k/k} + K_{k+1}\left(y_{k+1} - e^{-\lambda\delta}\widehat{\mu}_{k/k}\right), \tag{2.5}$$

$$\Gamma_{k+1/k+1} = (1 - K_{k+1})\,\Gamma_{k+1/k}, \tag{2.6}$$

with

$$\begin{aligned} K_{k+1} &= \frac{\Gamma_{k+1/k}}{\Gamma_{k+1/k} + \frac{\sigma_S^2}{\delta}}, \\ \Gamma_{k+1/k} &= e^{-2\lambda\delta}\Gamma_{k/k} + \frac{\sigma_\mu^2}{2\lambda}\left(1 - e^{-2\lambda\delta}\right). \end{aligned}$$

Stationary limit and continuous time representation

Solving the equation $\Gamma_{k+1/k+1} = \Gamma_{k/k}$ corresponding to the steady-state yields:

$$\begin{aligned} \Gamma_\infty &= \frac{g\left(\sigma_S, \lambda, \sigma_\mu\right) - f\left(\sigma_S, \lambda, \sigma_\mu\right)}{2e^{-2\lambda\delta}}, \\ \text{where } f\left(\sigma_S, \lambda, \sigma_\mu\right) &= \left(\frac{\sigma_S^2}{\delta} + \frac{\sigma_\mu^2}{2\lambda}\right)\left(1 - e^{-2\lambda\delta}\right), \\ \text{and } g\left(\sigma_S, \lambda, \sigma_\mu\right) &= \sqrt{f\left(\sigma_S, \lambda, \sigma_\mu\right)^2 + \frac{2\sigma_S^2\sigma_\mu^2}{\lambda\delta}\left(e^{-2\lambda\delta} - e^{-4\lambda\delta}\right)}. \end{aligned}$$

Using the stationary covariance error Γ_∞, a stationary gain K_∞ is defined and the estimate can be rewritten as a corrected exponential average:

$$\widehat{\mu}_{n+1} = K_\infty \sum_{i=0}^{\infty} e^{-\lambda\delta i}\left(1 - K_\infty\right)^i y_{n+1-i}. \tag{2.7}$$

The steady-state Kalman filter has also a continuous-time limit that depends on the asset returns. This result is recalled in the following proposition:

Proposition 13. *The steady-state Kalman filter $\widehat{\mu}$ solves the following stochastic differential equation*

$$d\widehat{\mu}_t = -\lambda\beta\left(\lambda, \sigma_\mu, \sigma_S\right)\widehat{\mu}_t dt + \lambda\left(\beta\left(\lambda, \sigma_\mu, \sigma_S\right) - 1\right)\frac{dS_t}{S_t}, \tag{2.8}$$

where

$$\beta\left(\lambda, \sigma_\mu, \sigma_S\right) = \left(1 + \frac{\sigma_\mu^2}{\lambda^2\sigma_S^2}\right)^{\frac{1}{2}}. \tag{2.9}$$

Proof. Based on Lakner (1998), the Kalman filter is given by:

$$\begin{aligned} E\left[\mu_t|\mathcal{F}_t^S\right] &= \phi(t)\left(\widehat{\mu}_0 + \frac{1}{\sigma_S^2}\int_0^t \frac{P(u)}{\phi(u)}\frac{dS_u}{S_u}\right), \\ \phi(t) &= e^{-\lambda t - \frac{1}{\sigma_S^2}\int_0^t P(u)du}, \end{aligned}$$

where the estimation error variance P is the solution of the following Riccati equation:

$$P'(t) = \frac{-1}{\sigma_S^2}P(t)^2 - 2\lambda P(t) + \sigma_\mu^2.$$

In this steady-state regime, we have $P'(t) = 0$. Then, the positive solution of this equation is given by

$$P^\infty = \sigma_S^2\lambda\left(\beta(\lambda,\sigma_\mu,\sigma_S) - 1\right),$$

and there holds:

$$\begin{aligned} \widehat{\mu}_t &= \phi^\infty(t)\left(\widehat{\mu}_0 + \frac{1}{\sigma_S^2}\int_0^t \frac{P^\infty}{\phi^\infty(u)}\frac{dS_u}{S_u}\right), \\ \phi^\infty(t) &= e^{-\lambda\beta(\lambda,\sigma_\mu,\sigma_S)t}. \end{aligned}$$

Since:

$$\frac{d\phi^\infty(t)}{\phi^\infty(t)} = -\lambda\beta(\lambda,\sigma_\mu,\sigma_S)\,dt,$$

Equation (2.8) follows. □

The Kalman filter is the optimal estimator for linear systems with Gaussian uncertainty, and such a continuous-time representation can be used for risk/return analysis of trend following strategies (see Bruder & Gaussel (2011) for details). Of course, in practice, the parameters $\theta = (\lambda, \sigma_\mu)$ are unknown and must be estimated. This important question is addressed in the next section.

2.2 Inference of the trend parameters

In this section, the problem of parameters inference is tackled, based on the use of statistical estimators such as Maximum Likelihood or Bayesian estimators. Two on-line computations of the likelihood are presented. Then, using the results of Genon-Catalot (see Genon-Catalot (2009) for details), we analyze the asymptotic behaviours of statistical estimators and provide the Cramer-Rao bound in closed form.

2.2.1 Likelihood computation

The likelihood can be computed using two methods. The first one is based on a direct calculus while the second method uses the Kalman filter.

Direct computation of the likelihood

A first approach is to directly compute the likelihood. The vectorial representation of the discrete time model (2.3)-(2.4) is:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_N \end{pmatrix},$$

where $(\mu_1, \cdots, \mu_N)^T$ and $(u_1, \cdots, u_N)^T$, knowing $\theta = (\sigma_\mu, \lambda)$, are two independent Gaussian processes. Therefore the vector $(y_1, \cdots, y_N)^T$, knowing θ, is also a Gaussian process. The likelihood is then characterized by the mean $M_{y_{1:N}|\theta}$ and the covariance $\Sigma_{y_{1:N}|\theta}$:

$$\begin{aligned} M_{y_{1:N}|\theta} &= 0 \ (\mu_0 = 0 \text{ is assumed}), & (2.10)\\ \Sigma_{y_{1:N}|\theta} &= \Sigma_{\mu_{1:N}|\theta} + \Sigma_{u_{1:N}|\theta}, & (2.11) \end{aligned}$$

where $\Sigma_{u_{1:N}|\theta} = \frac{\sigma_S^2}{\delta} I_N$ and $\Sigma_{\mu_{1:N}|\theta} = \left(\mathbb{C}ov\left(\mu_t, \mu_s\right)\right)_{1 \le t,s \le N}$. Since the drift μ is an Ornstein Uhlenbeck process, then:

$$\mathbb{C}ov\left(\mu_t, \mu_s\right) = \frac{\sigma_\mu^2}{2\lambda} e^{-\lambda(s+t)} \left(e^{2\lambda s \wedge t} - 1\right). \qquad (2.12)$$

Finally, the likelihood is given by:

$$f\left(y_1, ...y_N|\theta\right) = \frac{1}{(2\pi)^{N/2}\sqrt{det\Sigma_{y_{1:N}|\theta}}} e^{\left(\frac{-1}{2}(y_1,...,y_N)\Sigma_{y_{1:N}|\theta}^{-1}(y_1,...,y_N)^T\right)}. \qquad (2.13)$$

Remark 2.2.1. *When the dimension N is large, it is difficult and numerically unstable to directly invert the covariance matrix $\Sigma_{y_{1:N}|\theta}$ and compute its determinant. An iterative approach, the details of which are given in Appendix B, is preferred.*

Computation of the likelihood using the Kalman filter

The likelihood can also be evaluated via the prediction error decomposition (see Schweppe (1965) for details):

$$\begin{aligned} f\left(y_1, ...y_N|\theta\right) &= f\left(y_N|y_1, ...y_{N-1}, \theta\right) f\left(y_1, ...y_{N-1}|\theta\right). \\ &= \prod_{n=1}^{N} f\left(y_n|y_1, ...y_{n-1}, \theta\right), \end{aligned}$$

where the conditional laws are given in the following proposition:

Proposition 14. *The process* $(y_n|y_1,...y_{n-1},\theta)$ *is gaussian:*

$$(y_n|y_1,...y_{n-1},\theta) \quad \sim \quad \mathcal{N}\left(M_{y_{n|n-1}}, \mathbb{V}ar_{y_{n|n-1}}\right),$$

with

$$\begin{aligned} M_{y_{n|n-1}} &= e^{-\lambda\delta}\hat{\mu}_{n-1/n-1}, \\ \mathbb{V}ar_{y_{n|n-1}} &= e^{-2\lambda\delta}\Gamma_{n-1/n-1} + \frac{\sigma_\mu^2}{2\lambda}\left(1-e^{-2\lambda\delta}\right) + \frac{\sigma_S^2}{\delta}. \end{aligned}$$

The a posteriori estimate of the trend $\hat{\mu}_{n-1/n-1}$ *and the covariance error* $\Gamma_{n-1/n-1}$ *are given by Kalman filtering (see Equations (2.5) and (2.6)).*

Proof. Since the process y_n is gaussian, so is the process $(y_n|y_1,...y_{n-1},\theta)$. Moreover, using Equations (2.3)-(2.4), we have:

$$\begin{aligned} M_{y_{n|n-1}} &= \hat{\mu}_{n/n-1} + 0, \\ \hat{\mu}_{n/n-1} &= e^{-\lambda\delta}\hat{\mu}_{n-1/n-1} + 0, \end{aligned}$$

and

$$\begin{aligned} \mathbb{V}ar_{y_{n|n-1}} &= \Gamma_{n/n-1} + \frac{\sigma_S^2}{\delta}, \\ \Gamma_{n/n-1} &= e^{-2\lambda\delta}\Gamma_{n-1/n-1} + \frac{\sigma_\mu^2}{2\lambda}\left(1-e^{-2\lambda\delta}\right). \end{aligned}$$

□

Remark 2.2.2. *In practice, the volatility is not constant. However, if the volatility* σ_S *is* $\mathbf{F}^S$*-adapted, the two methods can be adapted and implemented. This assumption is satisfied if the volatility is a continuous process.*

2.2.2 Performance of statistical estimators

In this subsection, the asymptotic behaviour of the classical estimators is investigated.

Asymptotic behaviour of statistical estimator

The discrete time model (2.3)-(2.4) can be reformulated using the following proposition (see Genon-Catalot (2009) for details):

Proposition 15. *Consider the model* (2.3)-(2.4) *with* $(\lambda,\sigma_\mu,\sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$. *In this case, the process* (y_i) *is ARMA*$(1,1)$.

The asymptotic behaviour of the classical estimators follows. Indeed, the identifiability property and the asymptotic normality of the maximum likelihood estimator are well known for stationary ARMA gaussian processes (see Brockwell & Davis (2002b), section 10.8). Moreover, the asymptotic behaviour of the Bayesian estimators are also guaranteed by the ARMA$(1,1)$ property of the process (y_i). If the prior density function is continuous and positive in an open neighbourhood of the real parameters, the Bayesian estimators are asymptotically normal (see Tamaki (2008) in which a generalized Bernstein-Von Mises theorem for stationary "short memory" processes is given, or Monahan (1983) for a discussion on the Bayesian analysis of ARMA processes).

Cramer-Rao bound

This bound is the lowest variance of the unbiased estimators. We recall the following result, providing a formal description of the Cramer-Rao bound (CRB in short).

Corollary 2.2.3. *Consider the model* (2.3)-(2.4) *and* N *observations* $(y_1, \cdots, y_N)^T$. *Suppose that* $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$. *If* $\hat{\theta}_N$ *is an unbiased estimator of* $\theta = (\lambda, \sigma_\mu)$, *we have:*

$$\mathbb{C}ov_\theta \left(\hat{\theta}_N \right) \geqslant CRB\left(\theta\right).$$

This bound is given by $CRB\left(\theta\right) = I_N^{-1}\left(\theta\right)$, *where* $I_N\left(\theta\right)$ *is the Fisher Information matrix:*

$$\left(I_N\left(\theta\right)\right)_{i,j} = -\mathbb{E}\left[\frac{\partial^2 \log f\left(y_1, ...y_N | \theta\right)}{\partial\theta_i \partial\theta_j}\right],$$

and $I_N\left(\theta\right) = N I_1\left(\theta\right)$. *Moreover, the maximum likelihood estimator* $\widehat{\theta}_N^{ML}$ *attains this bound:*

$$\sqrt{N}\left(\widehat{\theta}_N^{ML} - \theta\right) \to \mathcal{N}\left(0, I_1^{-1}\left(\theta\right)\right).$$

This result is a consequence of Proposition 15 (see Brockwell & Davis (2002b), section 10.8). The following result is an analytic representation of the Fisher information matrix:

Theorem 2.2.4. *For the model* (2.3)-(2.4), *if* $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$, *we have:*

$$I_1\left(\theta\right) = \left(\frac{1}{4\Pi}\int_{-\Pi}^{\Pi} f_\theta^{-2}\left(\omega\right) \frac{\partial f_\theta}{\partial\theta_i}\left(\omega\right) \frac{\partial f_\theta}{\partial\theta_j}\left(\omega\right) d\omega\right)_{1 \leq i,j \leq 2},$$

where f_θ *is the spectral density of the process* (y_i):

$$f_\theta\left(\omega\right) = \frac{\frac{\sigma_\mu^2}{2\lambda}\left(1 - e^{-2\lambda\delta}\right) + \frac{\sigma_S^2}{\delta}\left(1 + e^{-2\lambda_\mu\delta}\right) - \frac{2e^{-\lambda_\mu\delta}\sigma_S^2}{\delta}\cos\left(\omega\right)}{1 + e^{-2\lambda_\mu\delta} - 2e^{-\lambda_\mu\delta}\cos\left(\omega\right)}.$$

Proof. Whittle's formula (see Whittle (1953) for details) gives the integral representation of the Fisher information matrix. Since the process (y_i) is ARMA$(1,1)$, the expression of its spectral density follows (see Brockwell & Davis (2002b), section 4.4). $\square$

Finally, the Cramer-Rao Bound of the trend parameters can be computed using Theorem 2.2.4.

2.3 Impact of parameters misspecification

In this section, we consider the continuous-time Kalman filter with a bad calibration in the steady-state regime. The law of the residuals between the filter (mis-specified or not) and the hidden process is characterized, and the impact of parameters misspecification on the detection of a positive trend is studied.

2.3.1 Context

Suppose that the risky asset S is given by the model (2.1)-(2.2) with $\theta^* = \left(\sigma_\mu^*, \lambda^*\right)$, and suppose that an agent thinks that the parameters are equal to $\theta = (\sigma_\mu, \lambda)$. Assuming the steady-state regime and using these estimates and Proposition 13, the agent implements the continuous time misspecified Kalman filter:

$$d\widehat{\mu}_t = -\lambda\beta\widehat{\mu}_t dt + \lambda\left(\beta - 1\right)\frac{dS_t}{S_t}, \tag{2.14}$$

where $\beta = \beta\left(\lambda, \sigma_\mu, \sigma_S\right)$ (see Equation (2.9)) and $\widehat{\mu}_0 = 0$. The following lemma gives the law of the misspecified Kalman filter:

Lemma 2.3.1. *Consider the model* (2.1)-(2.2) *with* $\theta^* = \left(\sigma_\mu^*, \lambda^*\right)$. *The misspecified, continuous-time filter of Equation (2.14) is given by:*

$$\widehat{\mu}_t = \lambda\left(\beta - 1\right)e^{-\lambda\beta t}\left(\int_0^t e^{\lambda\beta s}\mu_s^* ds + \sigma_S\int_0^t e^{\lambda\beta s}dW_s^S\right). \tag{2.15}$$

Moreover, $\widehat{\mu}$ is a centered gaussian process and its variance is given by:

$$\begin{aligned}\mathbb{V}ar\left[\widehat{\mu}_t\right] = \mathbb{E}\left[\widehat{\mu}_t^2\right] = &\frac{\lambda^2\left(\beta-1\right)^2\left(\sigma_\mu^*\right)^2}{\lambda^*\left(\lambda\beta-\lambda^*\right)}\left[\frac{1-e^{-(\lambda\beta+\lambda^*)t}}{\lambda\beta+\lambda^*}\right.\\ &\left.+\frac{2e^{-(\lambda\beta+\lambda^*)t}-e^{-2\lambda^* t}-e^{-2\lambda\beta t}}{\lambda\beta-\lambda^*}+\frac{e^{-2\lambda\beta t}-1}{2\lambda\beta}\right]\\ &+\frac{\lambda\left(\beta-1\right)^2\sigma_S^2}{2\beta}\left(1-e^{-2\lambda\beta t}\right).\end{aligned}$$

Proof. Applying Itō's lemma to the function $f(\widehat{\mu}, t) = e^{\lambda\beta t}\widehat{\mu}_t$, and integrating from 0 to t yields Equation (2.15). Therefore, $\widehat{\mu}$ is also a gaussian process. Its mean is zero (because $\mu^*_0 = 0$). Since the processes μ^* and W^S are supposed to be independent, the variance of $\widehat{\mu}$ is given by the sum of the variances of the terms in Equation (2.15). Moreover:

$$\begin{aligned}
\mathbb{V}ar\left[\int_0^t e^{\lambda\beta} dW_s^S\right] &= \frac{e^{2\lambda\beta t}-1}{2\lambda\beta}, \\
\mathbb{V}ar\left[\int_0^t e^{\lambda\beta s}\mu^*_s ds\right] &= \int_0^t\int_0^t e^{\lambda\beta(s_1+s_2)}\mathbb{C}ov\left(\mu^*_{s_1}, \mu^*_{s_2}\right) ds_1 ds_2,
\end{aligned}$$

and $\mathbb{C}ov\left(\mu^*_{s_1}, \mu^*_{s_2}\right)$ is given by Equation (2.12). The variance of the process $\widehat{\mu}_t$ follows. □

2.3.2 Filtering with parameters misspecification

The impact of parameters misspecification on trend filtering can be measured using the difference between the filter and the hidden process.
The following theorem gives the law of the residuals.

Theorem 2.3.2. *Consider the model (2.1)-(2.2) with $\theta^* = \left(\sigma^*_\mu, \lambda^*\right)$ and the trend estimate defined in Equation (2.15). The process $\widehat{\mu} - \mu^*$ is a centered gaussian process and its variance has a stationary limit:*

$$\lim_{t\to\infty} \mathbb{V}ar\left[\widehat{\mu}_t - \mu^*_t\right] = \frac{\sigma_S^2}{2\beta}\left(\lambda(\beta-1)^2 + \lambda^*\left((\beta^*)^2 - 1\right)\frac{\lambda^*\beta+\lambda}{\lambda\beta+\lambda^*}\right), \quad (2.16)$$

where $\beta = \beta(\lambda, \sigma_\mu, \sigma_S)$ and $\beta^* = \beta\left(\lambda^*, \sigma^*_\mu, \sigma_S\right)$ as given in Equation (2.9). Moreover, if $(\sigma_\mu, \lambda) = \left(\sigma^*_\mu, \lambda^*\right)$, Equation (2.16) becomes:

$$\lim_{t\to\infty} \mathbb{V}ar\left[\widehat{\mu}^*_t - \mu^*_t\right] = \lambda^*\sigma_S^2(\beta^*-1). \quad (2.17)$$

Proof. Equation (2.15) implies that the process $\widehat{\mu} - \mu^*$ is a centered gaussian process. Its variance can be computed in closed form:

$$\mathbb{V}ar\left[\widehat{\mu}_t - \mu^*_t\right] = \mathbb{V}ar\left[\widehat{\mu}_t\right] + \mathbb{V}ar\left[\mu^*_t\right] - 2 * \mathbb{C}\mathrm{ov}\left[\widehat{\mu}_t, \mu^*_t\right],$$

where $\mathbb{V}ar\left[\mu^*_t\right] = \frac{\left(\sigma^*_\mu\right)^2}{2\lambda^*}\left(1-e^{-2\lambda^* t}\right)$, and $\mathbb{V}ar\left[\widehat{\mu}_t\right]$ is given by Lemma 2.3.1. Since the processes W^S and μ^* are supposed to be independent, there holds:

$$\begin{aligned}
\mathbb{C}\mathrm{ov}\left[\widehat{\mu}_t, \mu^*_t\right] &= \frac{\lambda(\beta-1)\left(\sigma^*_\mu\right)^2}{2\lambda^*}\left(\frac{1-e^{-(\lambda\beta+\lambda^*)t}}{\lambda\beta+\lambda^*}\right. \\
&\quad \left. - \frac{e^{-2\lambda^* t} - e^{-(\lambda\beta+\lambda^*)t}}{\lambda\beta-\lambda^*}\right).
\end{aligned}$$

The asymptotic variance is obtained by letting $t \to \infty$:

$$\lim_{t\to\infty} \mathbb{V}ar\left[\widehat{\mu}_t - \mu_t^*\right] = \frac{\lambda(\beta-1)}{2\beta}\left[(\beta-1)\,\sigma_S^2 - \frac{\left(\sigma_\mu^*\right)^2(\beta+1)}{\lambda^*(\lambda\beta+\lambda^*)}\right] + \frac{\left(\sigma_\mu^*\right)^2}{2\lambda^*},$$

and Equation (2.16) follows. Finally, Equation (2.17) is obtained by letting $\theta \to \theta^*$. □

Remark 2.3.3. *Consider the well-specified case* $(\sigma_\mu, \lambda) = \left(\sigma_\mu^*, \lambda^*\right)$. *Using Equation (2.17), it follows that:*

$$\lim_{t\to\infty} \frac{\mathbb{V}ar\left[\widehat{\mu}_t^* - \mu_t^*\right]}{\mathbb{V}ar\left[\mu_t^*\right]} = \frac{2}{1+\sqrt{1+\frac{\left(\sigma_\mu^*\right)^2}{\left(\lambda^*\right)^2\sigma_S^2}}}. \tag{2.18}$$

Then, the asymptotic relative variance of the well-specified residuals is an increasing function of λ^* *and a decreasing function of* σ_μ^*.

2.3.3 Detection of a positive trend

In practice, the trend estimate (misspecified or not) will be used to make an investment decision. For example, a positive estimate leads to a long position. So, it is interesting to estimate the probability of a **positive** trend knowing a **positive** estimate. We derive this probability in closed form, based on the following proposition giving the asymptotic conditional law of the trend $(\mu_t^*|\hat{\mu}_t = x)$:

Proposition 16. *Consider the model* (2.1)-(2.2) *with* $\theta^* = \left(\sigma_\mu^*, \lambda^*\right)$ *and the trend estimate defined in Equation (2.15). Then, there holds:*

$$(\mu_t^*|\hat{\mu}_t = x) \underset{t\to\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}\left(\mathbb{M}^\infty_{\mu^*|\hat{\mu}}, \mathbb{V}ar^\infty_{\mu^*|\hat{\mu}}\right), \tag{2.19}$$

with:

$$\mathbb{M}^\infty_{\mu^*|\hat{\mu}} = \frac{\lambda_\mu^*\beta\left((\beta^*)^2-1\right)}{(\beta-1)\left(\lambda_\mu\beta+\lambda_\mu^*(\beta^*)^2\right)}x, \tag{2.20}$$

$$\mathbb{V}ar^\infty_{\mu^*|\hat{\mu}} = \mathbb{V}ar^\infty_{\mu^*}\left(1-\frac{\lambda_\mu^*\lambda_\mu\beta\left((\beta^*)^2-1\right)}{\left(\lambda_\mu^*+\lambda_\mu\beta\right)\left(\lambda_\mu\beta+\lambda_\mu^*(\beta^*)^2\right)}\right), \tag{2.21}$$

where $\mathbb{V}ar^\infty_{\mu^*} = \frac{(\sigma_\mu^*)^2}{2\lambda_\mu^*}$.

Moreover, if $(\sigma_\mu, \lambda) = \left(\sigma^*_\mu, \lambda^*\right)$*, Equation (2.19) becomes:*

$$\left(\mu^*_t | \hat{\mu}^*_t = x\right) \underset{t\to\infty}{\overset{\mathcal{L}}{\to}} \mathcal{N}\left(x, \frac{2\mathbb{V}ar^\infty_{\mu^*}}{\beta^* + 1}\right), \tag{2.22}$$

where $\beta^* = \beta\left(\lambda^*, \sigma^*_\mu, \sigma_S\right)$ *(see Equation (2.9)).*

Proof. Since the estimate $\hat{\mu}$ and the trend μ^* are two centered and correlated gaussian processes (see Lemma 2.3.1 and the proof of Theorem 2.3.2), the conditional law $(\mu^*_t|\hat{\mu}_t = x)$ is Gaussian with a mean and a variance given by:

$$\begin{aligned}
\mathbb{M}_{\mu^*_t|\hat{\mu}_t} &= \frac{\mathbb{C}ov\left(\hat{\mu}_t, \mu^*_t\right)}{\mathbb{V}ar\left[\hat{\mu}_t\right]}x, \\
\mathbb{V}ar_{\mu^*_t|\hat{\mu}_t} &= \mathbb{V}ar\left[\mu^*_t\right] - \frac{\mathbb{C}ov\left(\hat{\mu}_t, \mu^*_t\right)^2}{\mathbb{V}ar\left[\hat{\mu}_t\right]}.
\end{aligned}$$

Using Lemma 2.3.1 and the expression of $\mathbb{C}ov\left(\hat{\mu}_t, \mu^*_t\right)$ in the proof of Theorem 2.3.2 yields

$$\begin{aligned}
\lim_{t\to\infty} \mathbb{M}_{\mu^*_t|\hat{\mu}_t} &= \mathbb{M}^\infty_{\mu^*|\hat{\mu}}, \\
\lim_{t\to\infty} \mathbb{V}ar_{\mu^*_t|\hat{\mu}_t} &= \mathbb{V}ar^\infty_{\mu^*|\hat{\mu}},
\end{aligned}$$

and Equation (2.19) follows. Finally, Equation (2.22) is obtained by letting $\theta \to \theta^*$. □

The following proposition is a consequence of the previous proposition. It gives the asymptotic probability to have a positive trend, knowing a positive estimate equal to x.

Proposition 17. *Consider the model* (2.1)-(2.2) *with* $\theta^* = \left(\sigma^*_\mu, \lambda^*\right)$ *and the trend estimate defined in Equation (2.15). In this case:*

$$\lim_{t\to\infty} \mathbb{P}\left(\mu^*_t > 0 | \hat{\mu}_t = x\right) = \mathbb{P}_\infty\left(\mu^* > 0 | \hat{\mu} = x\right), \tag{2.23}$$

where

$$\mathbb{P}_\infty\left(\mu^* > 0 | \hat{\mu} = x\right) = 1 - \Phi\left(\frac{-\mathbb{M}^\infty_{\mu^*|\hat{\mu}=x}}{\sqrt{\mathbb{V}ar^\infty_{\mu^*|\hat{\mu}=x}}}\right), \tag{2.24}$$

where $\mathbb{M}^\infty_{\mu^*|\hat{\mu}=x}$ *and* $\mathbb{V}ar^\infty_{\mu^*|\hat{\mu}=x}$ *are defined in Equations (2.20) and (2.21), and* Φ *is the cumulative distribution function of the standard normal law. Moreover, if* $x > 0$ *and* $(\sigma_\mu, \lambda) = \left(\sigma^*_\mu, \lambda^*\right)$*, this asymptotic probability becomes an increasing function of* σ^*_μ *and a decreasing function of* λ^*_μ*.*

Proof. Equations (2.23) and (2.24) follow from Proposition 16. Now, consider the well-specified case $(\sigma_\mu, \lambda) = \left(\sigma_\mu^*, \lambda^*\right)$ and $x > 0$. Using Equation (2.22), it follows that:

$$\mathbb{V}ar^{\infty}_{\mu^*|\hat{\mu}^*=x} = f\left(\sigma_\mu^*, \lambda^*, \sigma_S\right),$$

where

$$f\left(\sigma_\mu^*, \lambda^*, \sigma_S\right) = \frac{\left(\sigma_\mu^*\right)^2}{\lambda^*\left(1+\sqrt{1+\frac{\left(\sigma_\mu^*\right)^2}{\sigma_S^2(\lambda^*)^2}}\right)}.$$

Since

$$\begin{aligned}
\frac{\partial f}{\partial \lambda^*}\left(\sigma_\mu^*, \lambda^*, \sigma_S\right) &= \frac{-\left(\sigma_\mu^*\right)^2}{(\lambda^*)^2\left(1+\sqrt{1+\frac{\left(\sigma_\mu^*\right)^2}{\sigma_S^2(\lambda^*)^2}}+\frac{\left(\sigma_\mu^*\right)^2}{\sigma_S^2}\right)} \leq 0,\\
\frac{\partial f}{\partial \sigma_\mu^*}\left(\sigma_\mu^*, \lambda^*, \sigma_S\right) &= \frac{\lambda^*\sigma_\mu^*\sigma_S^2\sqrt{1+\frac{\left(\sigma_\mu^*\right)^2}{\sigma_S^2(\lambda^*)^2}}}{\left(\sigma_\mu^*\right)^2+\sigma_S^2\left(\lambda^*\right)^2} \geq 0,
\end{aligned}$$

the asymptotic well-specified probability to have a positive trend, knowing a positive estimate equal to x is an increasing function of σ_μ^* and a decreasing function of λ_μ^*. $\square$

Remark 2.3.4. *This probability is an increasing function of x. Indeed, it is easier to detect the sign of the real trend with a high estimate than with a low estimate. Moreover, this probability is always superior to 0.5. This is due to the non-zero correlation between the trend and the filter. As shown in the previous sections, trend filtering is easier with a small spot volatility. Here, the probability to make a good detection is also a decreasing function of the σ_S.*

2.4 Simulations

In this section, numerical simulations are performed, in order to make the reader aware of the trend filtering problem. First, the feasibility of trend forecasting with statistical estimator is illustrated on different trend regimes. Then, the effects of a bad forecast on trend filtering and on the detection of a positive trend are discussed.

2.4.1 Feasibility of trend forecasting

Suppose that only discrete-time observations are available and that the time step is equal to $\delta = 1/252$. In this case, the agent uses the daily returns of the risky asset to calibrate the trend. We also assume that the agent uses an unbiased estimator. Given T years of observations, The Cramer-Rao bound is given by:

$$CRB_T(\theta) = \frac{I_1^{-1}(\theta)}{T * 252},$$

where $I_1(\theta)$ is given by Theorem 2.2.4. The smallest confidence region is obtained with this matrix. In practice, the real values of the parameters θ are unknown and asymptotic confidence regions are computed (replacing θ by the estimates $\hat{\theta}$ in the Fisher information matrix $I_1\left(\hat{\theta}\right)$). Since the goal of this subsection is to evaluated the feasibility of this estimation problem, we suppose that we know the real values of the parameters. In such a case, the exact Cramer-Rao bound can be computed. Suppose that a target standard deviation x_i is fixed for the parameter θ_i. In this case, to reach the precision x_i, the length of the observations must be superior to:

$$T_i^x = \frac{\left(I_1^{-1}(\theta)\right)_{ii}}{252 * x_i^2}.$$

We consider a fixed spot volatility $\sigma_S = 30\%$, two target precisions for each parameter θ_i and we compute T_i^x for several configurations. Figures 2.1, 2.2, 2.3 and 2.4 represent the results. It is well-known that for a high measurement noise, which means a high spot volatility, the problem is harder because of a low signal-to-noise ratio. The higher the volatility, the longer the observations must be. Here, we observe that with a higher drift volatility σ_μ and a lower λ, the problem is easier. Indeed, the drift takes higher values and is more detectable. Moreover, the simulations show that the classical estimators are not adapted to such a weak signal-to-noise ratio: even after a long period of observations, the estimators exhibit high variances. Indeed, the shortest observation period is longer than 29 years. It corresponds to a target standard deviation equal to 0.5 for a parameter $\lambda_\mu = 1$, and a trend standard deviation equal to $\sigma_\mu (2\lambda_\mu)^{-1/2} \approx 63\%$. Therefore, for this configuration, after 30 years of observations, the standard deviation is equal to 50% of the real parameter value λ_μ. After 742 years, this standard deviation is equal to 10%. Even with this kind of regime, the trend forecast with a good precision is impossible.

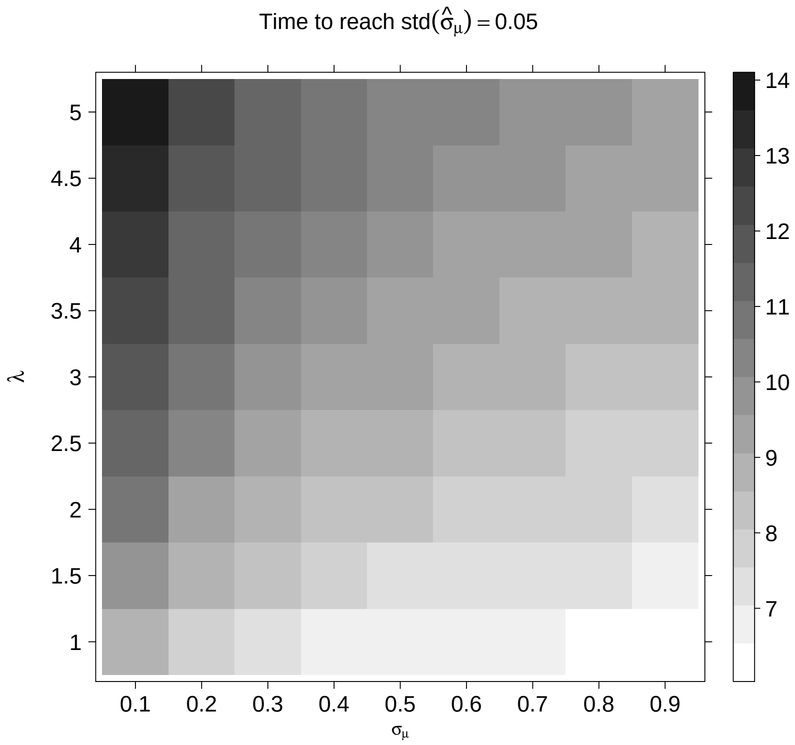

Figure 2.1: Time to reach a target standard deviation on σ_μ equal to 0.05 (ln(years))

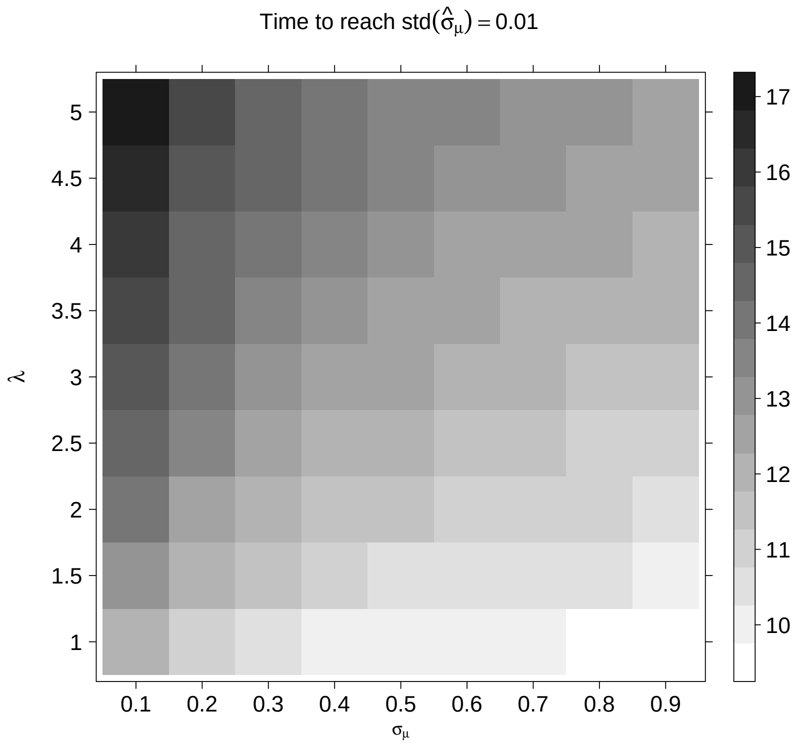

Figure 2.2: Time to reach a target standard deviation on σ_μ equal to 0.01 (ln(years))

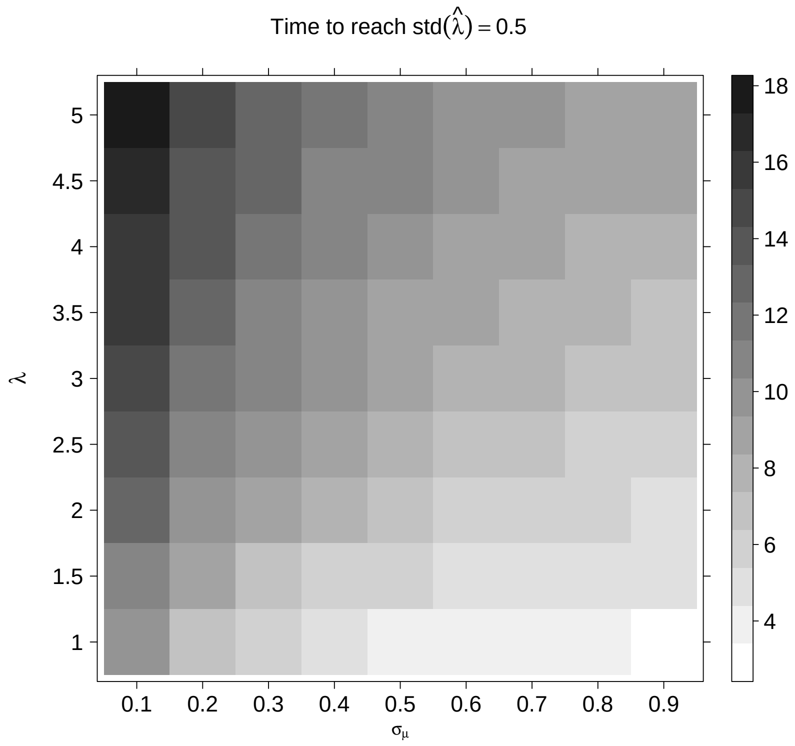

Figure 2.3: Time to reach a target standard deviation on λ equal to 0.5 (ln(years))

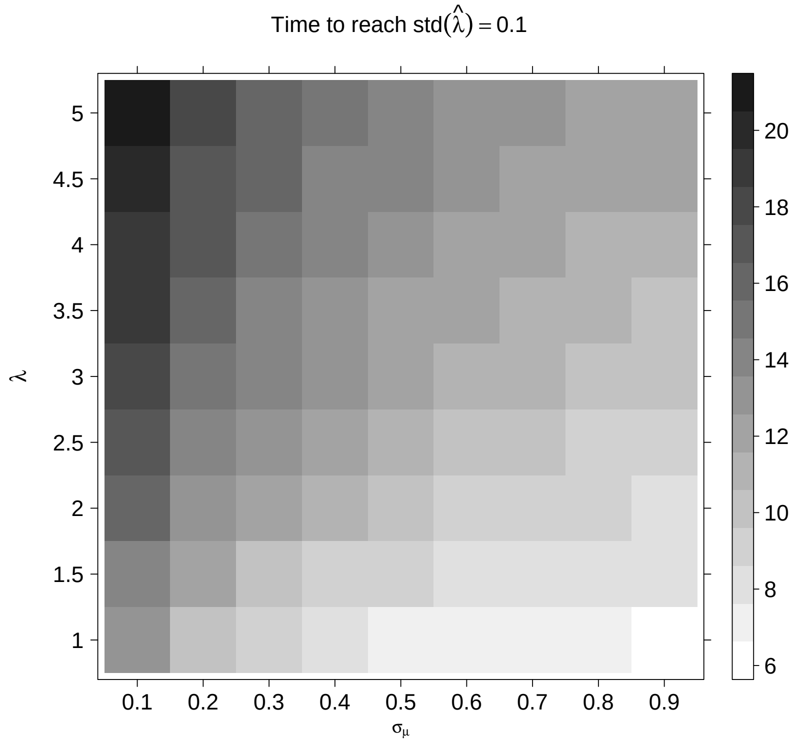

Figure 2.4: Time to reach a target standard deviation on λ equal to 0.1 (ln(years))

2.4.2 Impact of parameters misspecification on trend filtering

This subsection illustrates the impact of parameters misspecification on trend filtering. Using the results of Theorem 2.3.2, we represent, for different configurations, and for the well- and misspecified case, the asymptotic standard deviation of the residuals between the trend and the filter. Figures 2.5 and 2.6 represent the asymptotic standard deviation of the trend and of the residuals in the well-specified case (the agent uses the real values of the parameters) for different configurations. As seen in Equation (2.17), the asymptotic standard deviation of the well-specified residuals is an increasing function of the drift volatility σ_μ^* and a decreasing function of the parameter λ^*. For $\lambda^* = 1$ and $\sigma_\mu^* = 90\%$, the standard deviation of the residuals ($\simeq$ 44%) is inferior to the standard deviation of the trend ($\simeq$ 64%). For a high λ^* and a small drift volatility, the two quantities are approximately equal. This figure leads to the same conclusions than Equation (2.18). Indeed, as for the calibration problem, the problem of trend filtering is easier

with a small λ^* and a high drift volatility σ_μ^*.
Now consider the worst configuration $\sigma_S = 30\%$, $\lambda^* = 5$ and $\sigma_\mu^* = 10\%$. Figure 2.7 represents the asymptotic standard deviation of the residuals for different estimates (λ, σ_μ). This regime corresponds to a standard deviation of the trend equal to $\sigma_\mu^* \left(2\lambda_\mu^*\right)^{-1/2} \approx 3.2\%$ and to a standard deviation of the residuals equal to 3.16% in the well-specified case. If the agent implements the Kalman filter with $\lambda = 1$ and $\sigma_\mu = 90\%$, the standard deviation of the residuals becomes superior to 25%. Finally, consider the best configuration $\sigma_S = 30\%$, $\lambda^* = 1$ and $\sigma_\mu^* = 90\%$. Figure 2.8 represents the asymptotic standard deviation of the residuals for different estimates (λ, σ_μ). This regime corresponds to a trend standard deviation equal to $\sigma_\mu^* \left(2\lambda_\mu^*\right)^{-1/2} \approx 63\%$ and to a standard deviation of the residuals equal to 44% in the well-specified case. If the agent implements the Kalman filter with $\lambda = 5$ and $\sigma_\mu = 10\%$, the standard deviation of the residuals becomes superior to 60%. Even with a good regime, the impact of parameters mis-specification on trend filtering is not negligible.

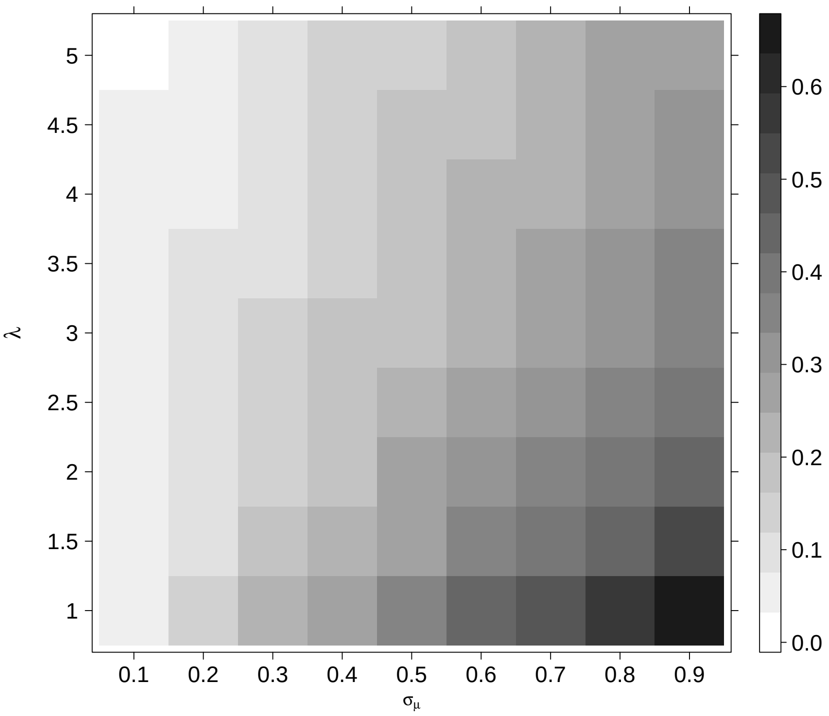

Figure 2.5: Asymptotic standard deviation of the trend as a function of the trend parameters with $\sigma_S = 30\%$.

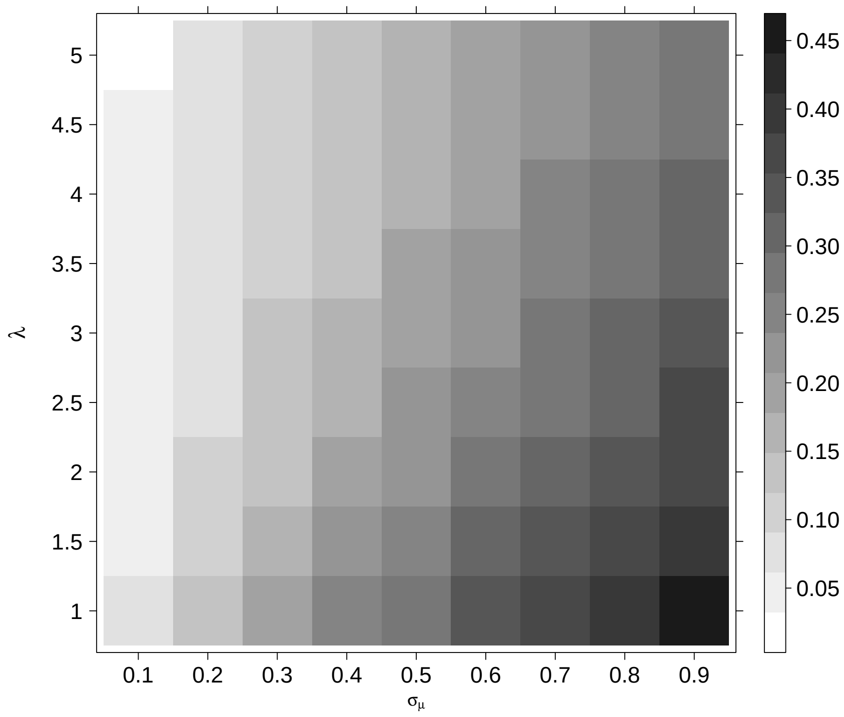

Figure 2.6: Asymptotic standard deviation of the residuals of the well-specified Kalman filter as a function of the trend parameters with $\sigma_S = 30\%$.

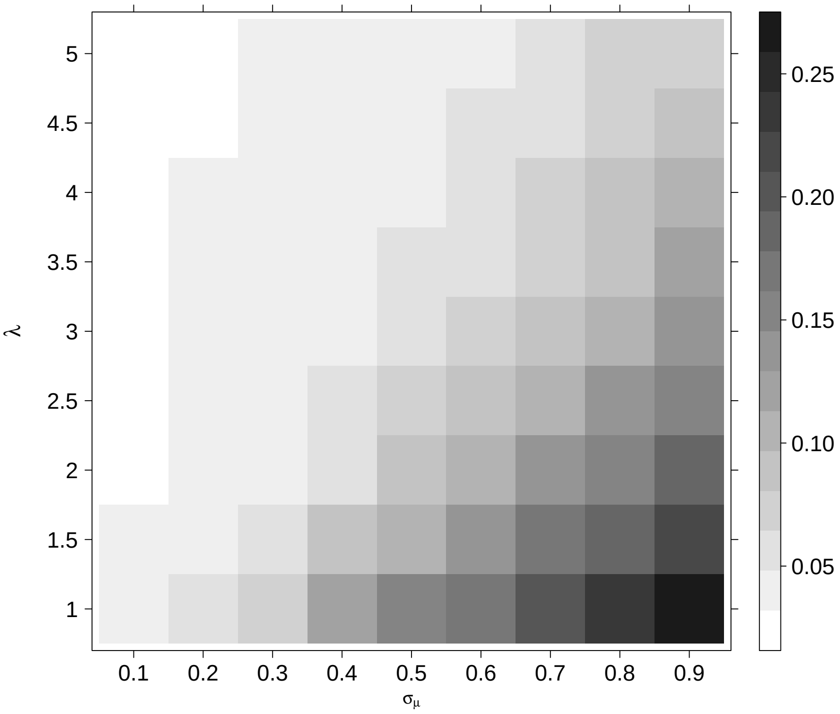

Figure 2.7: Asymptotic standard deviation of the residuals of the misspecified Kalman filter as a function of the trend estimate parameters with $\sigma_S = 30\%$, $\lambda^* = 5$ and $\sigma_\mu^* = 10\%$.

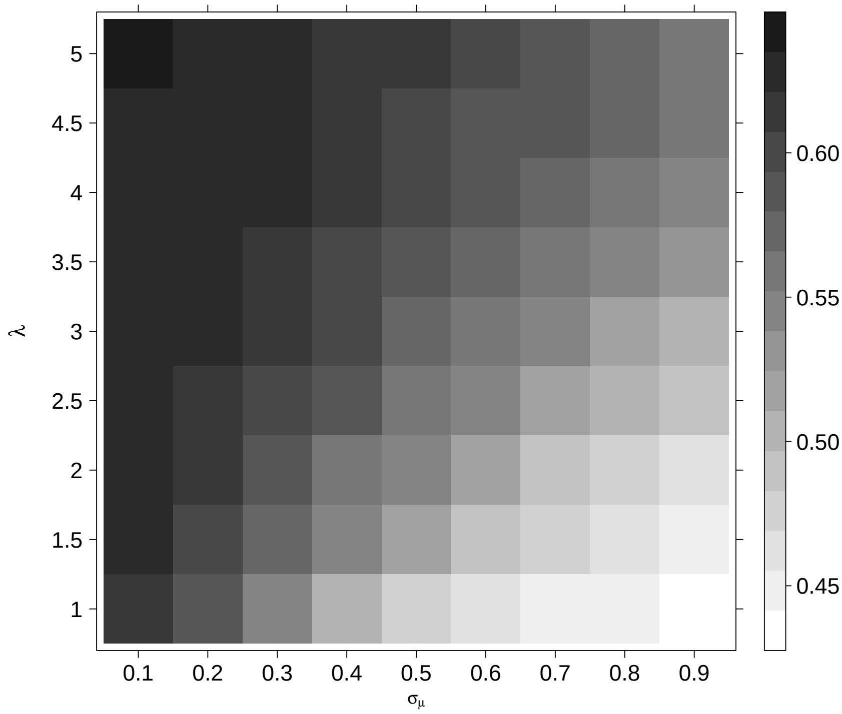

Figure 2.8: Asymptotic standard deviation of the residuals of the mis-specified Kalman filter as a function of the trend estimate parameters with $\sigma_S = 30\%$, $\lambda^* = 1$ and $\sigma_\mu^* = 90\%$.

2.4.3 Detection of a positive trend

In this subsection, Equation (2.24) - giving the asymptotic probability to have a positive trend, knowing a trend estimate equal to a threshold x - is illustrated.

In order to compare this probability for different trend regimes, we choose a threshold equal to the standard deviation of the filter $\hat{\mu}$. First, this quantity is tractable in practice. Moreover, since the continuous time mis-specified filter $\hat{\mu}$ is a centered Gaussian process, the probability that $\hat{\mu}$ becomes superior (or inferior) to its standard deviation is independent of the parameters $\left(\sigma_\mu^*, \lambda_\mu^*, \sigma_\mu, \lambda_\mu, \sigma_S\right)$. First, suppose that the agent uses the real values of the parameters and consider the asymptotic probability $\mathbb{P}\left(\mu^* > 0 | \hat{\mu}^* = \sqrt{\mathbb{V}_{\hat{\mu}^*}}\right)$ to have a positive trend, knowing an estimate equal to its standard deviation. Figure 2.9 represents this probability for different regimes. As seen in Proposition 17, in the well-specified case, this probability is an increasing function of the trend volatility σ_μ^* and a decreasing function of λ_μ^*. Again, as in the calibration and filtering problems, the detection is easier with a small λ^* and a high drift volatility. Now, suppose that the agent uses wrong estimates (σ_μ, λ). In this case, the agent implements the continuous time mis-specified Kalman filter. Figures 2.10 and 2.11 represent the asymptotic probability $\mathbb{P}\left(\mu^* > 0 | \hat{\mu} = \sqrt{\mathbb{V}_{\hat{\mu}}}\right)$ for the best and the worst configuration of Figure 2.9. As explained in Remark 2.3.4, this probability is always superior to 0.5, even with a bad calibration of the parameters. For each case, the probability to have a positive trend, knowing an estimate equal to its standard deviation does not vary a lot with an error on the parameters. This quantity seems to be robust to parameters mis-specifications.

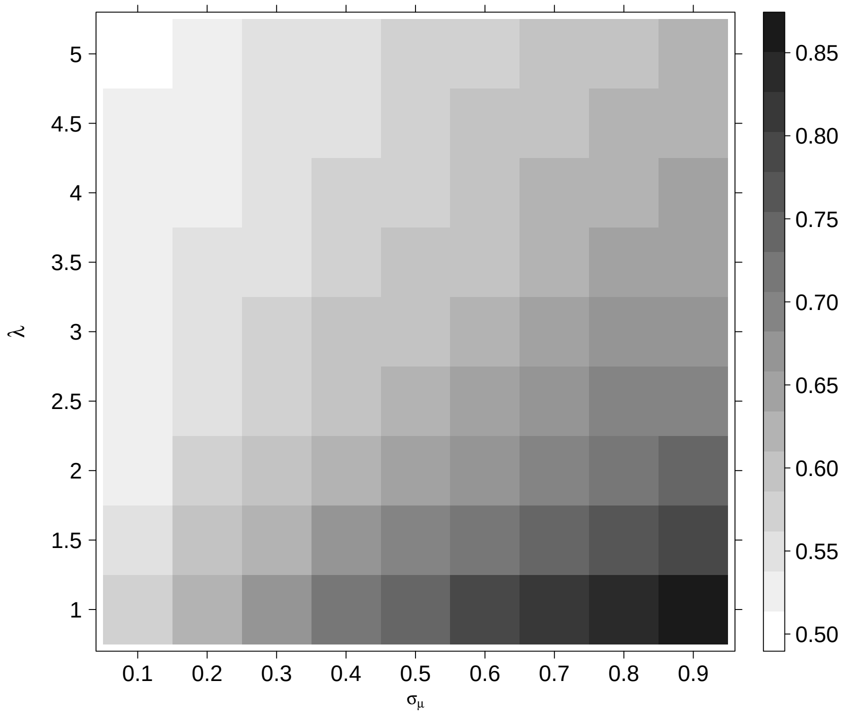

Figure 2.9: Asymptotic probability to have a positive trend given a well-specified estimate equal to its standard deviation with $\sigma_S = 30\%$.

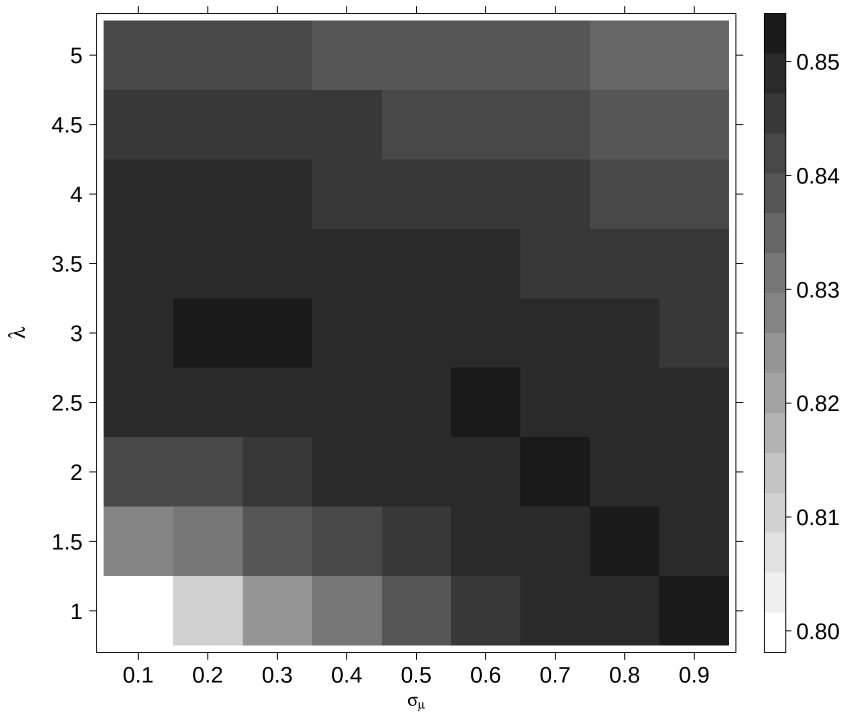

Figure 2.10: Asymptotic probability to have a positive trend given a misspecified estimate equal to its standard deviation with $\sigma_S = 30\%$, $\lambda^* = 1$ and $\sigma_\mu^* = 90\%$.

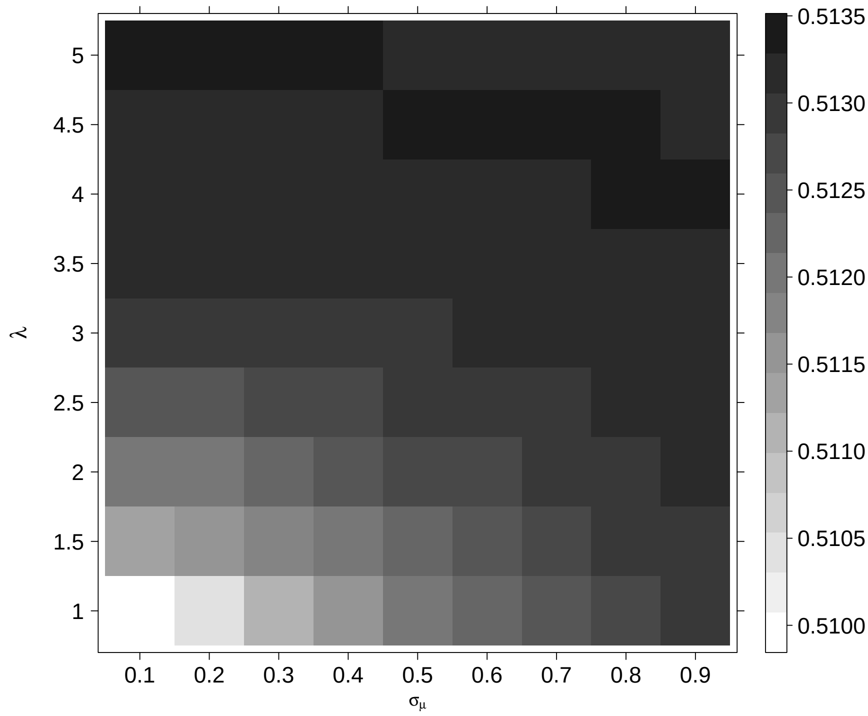

Figure 2.11: Asymptotic probability to have a positive trend given a misspecified estimate equal to its standard deviation with $\sigma_S = 30\%$, $\lambda^* = 5$ and $\sigma_\mu^* = 10\%$.

2.5 Conclusion

The present work tries to illustrate the difficulty of trend filtering with a model based on an unobserved mean-reverting diffusion. This model belongs to the class of Linear Gaussian Space State models. The advantage of this kind of system is to have an on-line method of estimation: the Kalman filter. In practice, the parameters of the model are unknown, and the calibration of filtering parameters becomes crucial. The linear and gaussian case allows to compute, in closed form, the likelihood. The Kalman filter can also be used for this analysis. These methods can be generalized to a non-constant volatility, and classical estimators can be easily put in practice.
Although this framework is particularly convenient for forecasting, the results of the analysis show that the classical estimators are not adapted to such a weak signal-to-noise ratio. The horizons of observations needed for a acceptable precision are too long. Therefore, the convergence is not guaranteed and the impact of mis-specification on trend filtering is not negligible. Despite these difficulties, the non-zero correlation between the trend and the estimate (mis-specified or not) can be used for the detection of a positive (or negative) trend.

Chapter 3

Performance analysis of the optimal strategy under partial information

The question addressed in this chapter is the performance of the optimal strategy, and the impact of partial information. The setting we consider is that of a stochastic asset price model where the trend follows an unobservable Ornstein-Uhlenbeck process. We focus on the optimal strategy with a logarithmic utility function under full or partial information. For both cases, we provide the asymptotic expectation and variance of the logarithmic return as functions of the signal-to-noise ratio and of the trend mean reversion speed. Finally, we compare the asymptotic Sharpe ratios of these strategies in order to quantify the loss of performance due to partial information.

Introduction

Optimal investment was introduced by Merton in 1969 (see Merton (1990) for details). He assumed that the risky asset follows a geometric Brownian motion and derived the optimal investment rules for an investor maximizing his expected utility function. Several generalisations of this problem are possible. One of them is to consider a stochastic unobservable trend, which leads to a system with partial information. This hypothesis seems to be realistic since only the historical prices of the risky asset are available to the public. For example, Karatzas & Zhao (2001) study the case of an unobservable constant trend, Lakner (1998) and Brendle (2006b) consider a stochastic asset price model where the trend is an unobservable Ornstein Uhlenbeck process, and Sass & Haussmann (2004) suppose that the trend is given by an unobserved continuous time, finite state Markov chain.
In this chapter, we consider a stochastic asset price model where the trend is an unobservable Ornstein Uhlenbeck process and we focus on the opti-

mal strategy with a logarithmic utility function under partial or complete information.
The purpose of this work is to characterize the performance of these strategies as functions of the signal-to-noise ratio and of the trend mean reversion speed and to quantify the loss of performance due to partial information. The loss of utility due to incomplete information was already studied in Karatzas & Zhao (2001), in Brendle (2006b) and in Rieder & Bauerle (2005). Here, the trading strategy performance is measured with the asymptotic Sharpe ratio (see Sharpe (1966) for details).
The chapter is organized as follows: the first section presents the model and recalls some results from filtering theory.
In the second section, the optimal strategy with complete information is investigated. This portfolio is built by an agent who is able to observe the trend and aims to maximize his expected logarithmic utility. We provide, in closed form, the expectation and variance of the logarithmic return as functions of the signal-to-noise ratio. We also show that the asymptotic Sharpe ratio of the optimal strategy with complete information is an increasing function of the signal-to-noise ratio.
In the third section, we consider the optimal strategy under partial information. This corresponds to an unobservable trend process and to an agent who aims to maximize his expected logarithmic utility. In this case, we provide, in closed form, the expectation and variance of the logarithmic return as functions of the signal-to-noise ratio and of the trend mean reversion speed. Then, we derive the asymptotic Sharpe ratio and we show that this is an increasing function of the signal-to-noise ratio and an unimodal (increasing then decreasing) function of the trend mean reversion speed. After that, we introduce the partial information factor which is the ratio between the asymptotic Sharpe ratio of the optimal strategy with partial information and the asymptotic Sharpe ratio of the optimal strategy with full information. This factor measures the loss of performance due to partial information. We show that this factor is bounded by a threshold equal to $\frac{2}{3^{3/2}}$.
In the fourth section, numerical examples illustrate the analytical results of the previous sections. The simulations show that, even with a high signal-to-noise ratio, a high trend mean reversion speed leads to a negligible performance of the optimal strategy under partial information compared to the case with complete information.

3.1 Setup

This section begins by presenting the model, which corresponds to an unobserved mean-reverting diffusion. After that, we reformulate this model in a completely observable environment (see Liptser & Shiriaev (1977) for details). This setting introduces the conditional expectation of the trend, knowing the past observations. Then, we recall the asymptotic continuous time limit of the Kalman filter.

3.1.1 The model

Consider a financial market living on a stochastic basis $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$, where $\mathbf{F} = \{\mathcal{F}_t, t \geqslant 0\}$ is the natural filtration associated to a two-dimensional (uncorrelated) Wiener process (W^S, W^μ), and $\mathbb{P}$ is the objective probability measure. The dynamics of the risky asset S is given by

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_S dW_t^S, \tag{3.1}$$

$$d\mu_t = -\lambda \mu_t dt + \sigma_\mu dW_t^\mu, \tag{3.2}$$

with $\mu_0 = 0$. We also assume that $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$. The parameter λ is called the trend mean reversion speed. Indeed, λ can be seen as the "force" that pulls the trend back to zero. Denote by $\mathbf{F}^S = \left\{\mathcal{F}_t^S\right\}$ be the natural filtration associated to the price process S. An important point is that only $\mathbf{F}^S$-adapted processes are observable, which implies that agents in this market do not observe the trend μ.

3.1.2 The observable framework

As stated above, the agents can only observe the stock price process S. Since, the trend μ is not F^S-measurable, the agents do not observe it directly. Indeed, the model (2.1)-(2.2) corresponds to a system with partial information. The following proposition gives a representation of the model (2.1)-(2.2) in an observable framework (see Liptser & Shiriaev (1977) for details or Appendix C for a proof).

Proposition 18. *The dynamics of the risky asset S is also given by*

$$\frac{dS_t}{S_t} = E\left[\mu_t | \mathcal{F}_t^S\right] dt + \sigma_S dN_t, \tag{3.3}$$

where N is a $\left(\mathbb{P}, \mathbf{F}^S\right)$ Wiener process.

Remark 3.1.1. *In the filtering theory (see Liptser & Shiriaev (1977) for details), the process N is called the innovation process. To understand this name, note that:*

$$dN_t = \frac{1}{\sigma_S}\left(\frac{dS_t}{S_t} - E\left[\mu_t | \mathcal{F}_t^S\right] dt\right).$$

Then, dN_t represents the difference between the current observation and what we expect knowing the past observations.

3.1.3 Optimal trend estimator

The system (2.1)-(2.2) corresponds to a Linear Gaussian Space State model (see Brockwell & Davis (2002a) for details). In this case, the Kalman filter gives the optimal estimator, which corresponds to the conditional expectation $E\left[\mu_t|\mathcal{F}_t^S\right]$. Since $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$, the model (2.1)-(2.2) is a controllable and observable time invariant system. In this case, it is well known that the estimation error variance converges to an unique constant value (see Kalman *et al.* (1962) for details). This corresponds to the steady-state Kalman filter. The following proposition (see the first chapter for a proof) gives a first continuous representation of the steady-state Kalman filter:

Proposition 19. *The steady-state Kalman filter has a continuous time limit depending on the asset returns:*

$$d\widehat{\mu}_t = -\lambda\beta\widehat{\mu}_t dt + \lambda\left(\beta - 1\right)\frac{dS_t}{S_t}, \tag{3.4}$$

where

$$\beta = \left(1 + \frac{\sigma_\mu^2}{\lambda^2\sigma_S^2}\right)^{\frac{1}{2}}. \tag{3.5}$$

The following proposition gives a second representation of the steady-state trend estimator $\hat{\mu}$:

Proposition 20. *Based on Equation (3.4), it follows that:*

$$d\hat{\mu}_t \quad = \quad -\lambda\hat{\mu}_t dt + \lambda\sigma_S\left(\beta - 1\right)dN_t. \tag{3.6}$$

Proof. Replacing $\frac{dS_t}{S_t}$ in Equations (3.4) by the expression of Equation (3.3), we find Equation (3.6). □

Remark 3.1.2. *It is well known that the Kalman estimator is a Gaussian process. Here, we find that the steady-state trend estimator $\hat{\mu}$ is an Ornstein Uhlenbeck process. In practice, the parameters $(\lambda, \sigma_\mu, \sigma_S)$ are unknown and must be estimated (see the first chapter where we assess the feasibility of forecasting trends modeled by an unobserved mean-reverting diffusion). In this chapter, we assume that the parameters are known.*

3.2 Optimal strategy under complete information

In this section, the optimal strategy under full information is investigated. This strategy is built by an agent who is able to observe the trend μ. Formally, it corresponds to the case $\mathbf{F}^S = \mathbf{F}$. Given this framework, we consider the optimal strategy with a logarithmic utility function. We provide, in closed form, the asymptotic expectation and variance of the logarithmic return, and the asymptotic Sharpe ratio of this strategy as functions of the signal-to-noise ratio.

3.2.1 Context

Consider the financial market defined in the first section with a risk free rate and without transaction costs. Let P^o be a self financing portfolio given by:

$$\begin{aligned} \frac{dP_t^o}{P_t^o} &= \omega_t^o \frac{dS_t}{S_t}, \\ P_0^o &= x, \end{aligned}$$

where ω_t^o is the fraction of wealth invested in the risky asset (also named the control variable). The agent aims to maximize his expected logarithmic utility on an admissible domain $\mathcal{A}^o$ for the allocation process. In this section, we assume that the agent is able to observe the trend μ. Formally, it means that $\mathcal{A}^o$ represents all the $\mathbf{F}$-progressive and measurable processes and the solution of this problem is given by:

$$\omega^* = \arg \sup_{\omega \in \mathcal{A}^o} \mathbb{E}\left[\ln\left(P_t^o\right) | P_0^o = x\right].$$

As is well known (see Lakner (1998) or Bjork *et al.* (2010) for examples), the solution of this problem is given by:

$$\frac{dP_t^o}{P_t^o} = \frac{\mu_t}{\sigma_S^2} \frac{dS_t}{S_t}, \tag{3.7}$$

$$P_0^o = x. \tag{3.8}$$

3.2.2 Performance analysis of the optimal strategy under complete information

The following proposition gives the stochastic differential equation of the portfolio P^o:

Proposition 21. *Consider the portfolio P^o given by Equation (3.7). In this case,*

$$d\ln(P_t^o) = \frac{\mu_t^2}{2\sigma_S^2} dt + \frac{\mu_t}{\sigma_S} dW_t^S. \tag{3.9}$$

Proof. Using Equation (3.7) and Itô's lemma on the process $\ln(P_t^o)$, the result follows. □

The asymptotic expected logarithmic return is the first indicator to assess the potential of a trading strategy. The second one can be the variance of the logarithmic return. This indicator can be useful as a measure of risk.

Moreover, let SR_T be the annualized Sharpe-ratio at time T of a portfolio (P_T) defined by:

$$\mathrm{SR}_T = \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{\sqrt{T\ \mathbb{V}\mathrm{ar}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}}. \tag{3.10}$$

This indicator measures the expected logarithmic return per unit of risk. The Sharpe ratio is a prime metric for an investment.

Remark 3.2.1. *This definition of the Sharpe ratio is different from the original one (see Sharpe (1966)). Here, this indicator is computed on logarithmic returns.*

The following theorem gives the asymptotic expectation, variance and Sharpe ratio of the logarithmic return:

Theorem 3.2.2. *Consider the portfolio given by Equation (3.7). In this case:*

$$\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P^o_T}{P^o_0}\right)\right]}{T} = \frac{SNR}{2}, \tag{3.11}$$

$$\lim_{T\to\infty} \frac{\mathbb{V}ar\left[\ln\left(\frac{P^o_T}{P^o_0}\right)\right]}{T} = SNR, \tag{3.12}$$

$$SR^o_\infty = \frac{\sqrt{SNR}}{2}. \tag{3.13}$$

where SNR is the signal-to-noise-ratio:

$$SNR = \frac{\sigma_\mu^2}{2\lambda\sigma_S^2}. \tag{3.14}$$

Proof. Integrating the expression of Proposition 21 from 0 to T and taking the expectation, it gives:

$$\mathbb{E}\left[\ln\left(\frac{P^o_T}{P^o_0}\right)\right] = \frac{1}{2\sigma_S^2}\int_0^T \mathbb{E}\left[\mu_t^2\right] dt + 0.$$

Since μ is an Ornstein-Uhlenbeck process:

$$\begin{aligned} \mathbb{E}\left[\mu_t\right] &= 0, \\ \mathbb{V}\mathrm{ar}\left[\mu_t\right] &= \sigma_\mu^2 \frac{1-e^{-2\lambda t}}{2\lambda}. \end{aligned}$$

Then, tending T to ∞, Equation (3.11) follows. Since the processes W^S and μ are supposed to be independent:

$$\mathbb{V}\mathrm{ar}\left[\ln\left(\frac{P^o_T}{P^o_0}\right)\right] = \frac{1}{4\sigma_S^4}\mathbb{V}\mathrm{ar}\left[\int_0^T \mu_t^2 dt\right] + \frac{1}{\sigma_S^2}\mathbb{V}\mathrm{ar}\left[\int_0^T \mu_t dW_t^S\right].$$

Since the process $\left(\int_0^T \mu_t dW_t^S\right)_{T\geq 0}$ is a martingale:

$$\mathbb{V}\mathrm{ar}\left[\int_0^T \mu_t dW_t^S\right] = \int_0^T \mathbb{E}\left[\mu_t^2\right] dt = \frac{\sigma_\mu^2}{2\lambda}\left(T + \frac{1-e^{-2\lambda T}}{2\lambda}\right),$$

Moreover, Isserlis' theorem (see Isserlis (1918) for details) gives:

$$\mathbb{V}\mathrm{ar}\left[\int_0^T \mu_t^2 dt\right] = 2\int_0^T \int_0^T \left(\mathbb{E}\left[\mu_s \mu_t\right]\right)^2 ds dt.$$

Since μ is an Ornstein Uhlenbeck:

$$\mathbb{V}\mathrm{ar}\left[\int_0^T \mu_t^2 dt\right] = \frac{\sigma_\mu^4 e^{-4\lambda T}\left(e^{4\lambda T}\left(4\lambda T - 5\right) + e^{2\lambda T}\left(8\lambda T + 4\right) + 1\right)}{8\lambda^4}.$$

Equation (3.12) follows. Finally, using the definition of the Sharpe ratio (see Equation (3.10)) and the results of Equations (3.11) and (3.12), Equation (3.13) follows. □

Theorem 3.2.2 shows that the asymptotic expectation and the asymptotic variance logarithmic return are linear functions of the signal-to-noise ratio and that the asymptotic Sharpe ratio is a linear function of the ratio between the asymptotic trend standard deviation and the volatility.

3.3 Optimal strategy under partial information

In this section, the Merton's problem under partial information is investigated. We consider the case of a logarithmic utility function. We provide, in closed form, the asymptotic expectation and variance of the logarithmic return, and the asymptotic Sharpe ratio of this strategy as functions of the signal-to-noise ratio and of the trend mean reversion speed. After that, we introduce the partial information factor which is the ratio between the asymptotic Sharpe ratio of the optimal strategy with partial information and the asymptotic Sharpe ratio of the optimal strategy with complete information. We close this section by showing that this factor is bounded by a threshold equal to $\frac{2}{3^{3/2}}$.

3.3.1 Context

Consider the financial market defined in the first section with a risk free rate and without transaction costs. Let P be a self financing portfolio given by:

$$\begin{aligned} \frac{dP_t}{P_t} &= \omega_t \frac{dS_t}{S_t}, \\ P_0 &= x, \end{aligned}$$

where ω_t is the fraction of wealth invested in the risky asset. The agent aims to maximize his expected logarithmic utility on an admissible domain $\mathcal{A}$ for the allocation process. In this section, we assume that the agent is not able to observe the trend μ. Formally, $\mathcal{A}$ represents all the $\mathbf{F}^S$-progressive and measurable processes and the problem is:

$$\omega^* = \arg\sup_{\omega \in \mathcal{A}} \mathbb{E}\left[\ln\left(P_t\right)|P_0 = x\right].$$

The solution of this problem is well known and easy to compute (see Lakner (1998) for example). Indeed, it has the following form:

$$\omega_t^* = \frac{E\left[\mu_t|\mathcal{F}_t^S\right]}{\sigma_S^2}.$$

Using the steady-state Kalman filter, the optimal portfolio is given by:

$$\frac{dP_t}{P_t} = \frac{\widehat{\mu}_t}{\sigma_S^2}\frac{dS_t}{S_t}, \tag{3.15}$$

$$P_0 = x, \tag{3.16}$$

where $\widehat{\mu}$ is given by Equation (3.4).

3.3.2 Performance analysis of the optimal strategy under partial information

The following proposition gives the stochastic differential equation of the portfolio:

Proposition 22. *The optimal portfolio process of Equation (3.15) follows the dynamics:*

$$d\ln(P_t) = \frac{1}{2\sigma_S^2\lambda\left(\beta - 1\right)}d\widehat{\mu}_t^2 + \left[\frac{\widehat{\mu}_t^2}{\sigma_S^2}\left(\frac{\beta}{\left(\beta - 1\right)} - \frac{1}{2}\right) - \frac{1}{2}\lambda\left(\beta - 1\right)\right]dt,$$

where β is given by Equation (3.5).

Proof. Equation (3.15) is equivalent to (by Itô's lemma):

$$d\ln(P_t) = \frac{\widehat{\mu}_t}{\sigma_S^2}\frac{dS_t}{S_t} - \frac{1}{2}\frac{\widehat{\mu}_t^2}{\sigma_S^2}dt.$$

Using Equation (3.4),

$$d\ln(P_t) = \frac{\widehat{\mu}_t}{\sigma_S^2}\frac{d\widehat{\mu}_t}{\lambda\left(\beta - 1\right)} + \frac{\widehat{\mu}_t^2}{\sigma_S^2}\frac{\beta}{\left(\beta - 1\right)}dt - \frac{1}{2}\frac{\widehat{\mu}_t^2}{\sigma_S^2}dt,$$

Itô's lemma on Equation (3.4) gives:

$$d\widehat{\mu}_t^2 = 2\widehat{\mu}_t d\widehat{\mu}_t + \lambda^2\left(\beta_{\sigma_\mu,\lambda,\sigma_S} - 1\right)^2\sigma_S^2 dt.$$

Using this equation, the dynamic of the logarithmic wealth follows. $\square$

Remark 3.3.1. *Proposition 22 shows that the returns of the optimal strategy with partial information can be broken down into two terms. The first one represents an option on the square of the realized returns (called Option profile). The second term is called the Trading Impact. These terms are introduced and discussed in Bruder & Gaussel (2011). The option profile at the time T is:*

$$\text{Option Profile}_T = \frac{1}{2\sigma_S^2} \frac{1}{\lambda(\beta-1)} (\widehat{\mu}_T^2 - \widehat{\mu}_0^2).$$

With the assumption of an initial trend estimate equal to 0*, the Option profile is always positive. The Trading Impact is a cumulated function of the trend estimate:*

$$\text{Trading impact}_T = \int_0^T \left[\frac{\widehat{\mu}_t^2}{\sigma_S^2} \left(\frac{\beta}{(\beta-1)} - \frac{1}{2} \right) - \frac{1}{2}\lambda(\beta-1) \right] dt.$$

When $T \to \infty$, it becomes the preponderant term. The Trading Impact is positive on the long term T if the drift estimate $\widehat{\mu}_t$ verifies:

$$\frac{1}{T}\int_0^T \widehat{\mu}_t^2 dt > \frac{\lambda\sigma_S^2(\beta-1)}{2\frac{\beta}{\beta-1} - 1}, \tag{3.17}$$

Equation (3.17) can be seen as a condition for the trend following strategy to generate profits in the long term.

The following theorem gives the asymptotic expectation, variance and Sharpe ratio of the logarithmic return:

Theorem 3.3.2. *Consider the portfolio given by Equation (3.15). In this case:*

$$\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = \frac{\lambda}{4}(\beta-1)^2, \tag{3.18}$$

$$\lim_{T\to\infty} \frac{\mathbb{V}ar\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = \frac{\lambda}{8}\left(\beta^2-1\right)^2, \tag{3.19}$$

$$\lim_{T\to\infty} SR_T = \sqrt{\frac{\lambda}{2}\frac{\beta-1}{\beta+1}}, \tag{3.20}$$

where β is given by Equation (3.5).

Proof. Based on Equation (3.6), $\hat{\mu}$ is an Ornstein-Uhlenbeck process:

$$\begin{aligned}
\mathbb{E}[\hat{\mu}_t] &= 0, \\
\mathbb{V}\mathrm{ar}[\hat{\mu}_t] &= (\lambda\sigma_S(\beta-1))^2 \frac{1-e^{-2\lambda t}}{2\lambda}.
\end{aligned}$$

Integrating the expression of Proposition 22 from 0 to T and taking the expectation, it gives:

$$\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right] = \frac{(\beta-1)}{4}\left(1-e^{-2\lambda T}\right) - \frac{\lambda(\beta-1)}{2}T$$
$$+\left(\frac{\beta}{\beta-1}-\frac{1}{2}\right)\left[\frac{\lambda}{2}(\beta-1)^2 T - \frac{\lambda}{2}(\beta-1)^2\frac{1-e^{-2\lambda T}}{2\lambda}\right]$$

Then, tending T to ∞, Equation (3.18) follows. Integrating the expression of Proposition 22 from 0 to T and taking the variance, it gives:

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\ln\left(\frac{P_T}{P_0}\right)\right] &= \frac{1}{\left(2\lambda\left(\beta-1\right)\sigma_S^2\right)^2}\mathbb{V}\mathrm{ar}\left[\hat{\mu}_T^2\right] \\ &+\frac{1}{\sigma_S^4}\left(\frac{\beta}{\beta-1}-\frac{1}{2}\right)^2\mathbb{V}\mathrm{ar}\left[\int_0^T \hat{\mu}_t^2 dt\right] \\ &+\frac{2\left(\frac{\beta}{\beta-1}-\frac{1}{2}\right)}{\left(2\lambda\left(\beta-1\right)\sigma_S^4\right)}\mathbb{C}\mathrm{ov}\left[\hat{\mu}_T^2,\int_0^T \hat{\mu}_t^2 dt\right].\end{aligned}$$

Moreover

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\hat{\mu}_T^2\right] &= \mathbb{C}\mathrm{ov}\left[\hat{\mu}_T^2,\hat{\mu}_T^2\right], \\ \mathbb{V}\mathrm{ar}\left[\int_0^T \hat{\mu}_t^2 dt\right] &= \int_0^T\int_0^T \mathbb{C}\mathrm{ov}\left[\hat{\mu}_s^2,\hat{\mu}_t^2\right] dsdt, \\ \mathbb{C}\mathrm{ov}\left[\hat{\mu}_T^2,\int_0^T \hat{\mu}_t^2 dt\right] &= \int_0^T \mathbb{C}\mathrm{ov}\left[\hat{\mu}_s^2,\hat{\mu}_T^2\right] ds,\end{aligned}$$

and the expression of $\mathbb{C}\mathrm{ov}\left[\hat{\mu}_s^2,\hat{\mu}_t^2\right]$ is given in Lemma 4.5.3 (see Appendix D). Then

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left[\hat{\mu}_T^2\right] &= \frac{\lambda^2\sigma_S^4\left(\beta-1\right)^4}{2}\left(1-2e^{-2\lambda T}+e^{-4\lambda T}\right), \\ \mathbb{V}\mathrm{ar}\left[\int_0^T \hat{\mu}_t^2 dt\right] &= \frac{\lambda^2\sigma_S^4\left(\beta-1\right)^4}{2}\left(\frac{1-e^{-2\lambda T}}{2\lambda}\right. \\ &\left.+\frac{e^{-2\lambda T}-e^{-4\lambda T}}{2\lambda}-2Te^{-2\lambda T}\right), \\ \mathbb{C}\mathrm{ov}\left[\hat{\mu}_T^2,\int_0^T \hat{\mu}_t^2 dt\right] &= \frac{\lambda\sigma_S^4\left(\beta-1\right)^4}{2}\left(T-\frac{5}{4\lambda}+\frac{e^{-2\lambda T}}{\lambda}\right. \\ &\left.+\frac{e^{-4\lambda T}}{4\lambda T}+2Te^{-2\lambda T}\right).\end{aligned}$$

Finally, using these expressions and tending T to ∞, Equations (3.19) and (3.20) follow. □

The following result is a corollary of the previous theorem. It represents the asymptotic expectation, variance and Sharpe ratio of the logarithmic return as a function of the signal-to-noise-ratio and of the trend mean reversion speed λ.

Corollary 3.3.3. *Consider the portfolio given by Equation (3.15). In this case:*

$$\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = \frac{1}{2}\left(SNR + \lambda - \sqrt{\lambda\left(\lambda + 2SNR\right)}\right), \tag{3.21}$$

$$\lim_{T\to\infty} \frac{\mathbb{V}ar\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = \frac{SNR^2}{2\lambda}, \tag{3.22}$$

$$\lim_{T\to\infty} SR_T = \left(\frac{\lambda}{2}\right)^{3/2} \frac{\left(\sqrt{1+\frac{2SNR}{\lambda}} - 1\right)^2}{SNR}, \tag{3.23}$$

where SNR is the signal-to-noise-ratio (see Equation (3.14)). Moreover:

1. *For a fixed parameter value λ,*
   - *the asymptotic expected logarithmic return is an increasing function of SNR,*
   - *the asymptotic Sharpe ratio is an increasing function of SNR.*
2. *For a fixed parameter value SNR,*
   - *the asymptotic expected logarithmic return is a decreasing function of λ,*
   - *the asymptotic Sharpe ratio is a decreasing function of λ if:*

$$SNR < \frac{3}{2}\lambda, \tag{3.24}$$

and an increasing function of λ if $SNR > \frac{3}{2}\lambda$.

The maximum asymptotic Sharpe ratio is attained for $\lambda = \frac{2}{3}SNR$ and is equal to:

$$SR_{\infty}^{Max} = \frac{\sqrt{SNR}}{3^{3/2}}. \tag{3.25}$$

Proof. Using Equation (3.14) and Equation (3.5), it follows that:

$$\beta = \sqrt{1 + \frac{2\mathrm{SNR}}{\lambda}}.$$

Injecting this expression in Equation (3.18), we find:

$$\lim_{T \to \infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = L\left(\mathrm{SNR}, \lambda\right),$$

where

$$L\left(\mathrm{SNR}, \lambda\right) = \frac{1}{2}\left(\mathrm{SNR} + \lambda - \sqrt{\lambda\left(\lambda + 2\mathrm{SNR}\right)}\right).$$

Since

$$\frac{\partial L\left(\mathrm{SNR}, \lambda\right)}{\partial \mathrm{SNR}} = \frac{1}{2}\left(1 - \frac{1}{\sqrt{1 + \frac{2\mathrm{SNR}}{\lambda}}}\right) \geq 0,$$

the asymptotic expected logarithmic return is an increasing function of SNR. Moreover:

$$\frac{\partial L\left(\mathrm{SNR}, \lambda\right)}{\partial \lambda} = \frac{1}{2}\left(1 - \frac{\lambda + \mathrm{SNR}}{\sqrt{\lambda\left(\lambda + 2\mathrm{SNR}\right)}}\right) \leq 0,$$

it follows that the asymptotic expected logarithmic return is a decreasing function of λ. Moreover, using Equations (3.14), (3.5) and (3.19), Equation (3.22) follows.

Now, with Equations (3.14), (3.5) and (3.20), we find:

$$\lim_{T \to \infty} \mathrm{SR}_T = \mathrm{SR}_\infty\left(\mathrm{SNR}, \lambda\right),$$

where

$$\mathrm{SR}_\infty\left(\mathrm{SNR}, \lambda\right) = \left(\frac{\lambda}{2}\right)^{3/2} \frac{\left(\sqrt{1 + \frac{2\mathrm{SNR}}{\lambda}} - 1\right)^2}{\mathrm{SNR}}.$$

Since

$$\frac{\partial \mathrm{SR}_\infty}{\partial \mathrm{SNR}} = \frac{\lambda^{\frac{3}{2}}\left(\sqrt{1 + \frac{2\mathrm{SNR}}{\lambda}} - 1\right)^2}{2\mathrm{SNR}^2\sqrt{2\left(1 + \frac{2\mathrm{SNR}}{\lambda}\right)}} \geq 0,$$

the asymptotic Sharpe ratio is an increasing function of SNR. Moreover:

$$\frac{\partial \mathrm{SR}_\infty}{\partial \lambda} = \frac{\left(\sqrt{1 + \frac{2\mathrm{SNR}}{\lambda}} - 1\right)\left(3\lambda\left(1 - \sqrt{1 + \frac{2\mathrm{SNR}}{\lambda}}\right) + 2\mathrm{SNR}\right)}{4\sqrt{2\lambda}\mathrm{SNR}\sqrt{1 + \frac{2\mathrm{SNR}}{\lambda}}}.$$

Then, the sign of $\frac{\partial \mathrm{SR}_\infty}{\partial \lambda}$ is given by the sign of:

$$\mathrm{A}\left(\mathrm{SNR},\lambda\right) = \left(3\lambda\left(1-\sqrt{1+\frac{2\mathrm{SNR}}{\lambda}}\right)+2\mathrm{SNR}\right).$$

Using $\beta = \sqrt{1+\frac{2\mathrm{SNR}}{\lambda}}$, this expression can be factorised:

$$\mathrm{A}\left(\mathrm{SNR},\lambda\right) = \lambda\left(\beta-1\right)\left(\beta-2\right).$$

Since $\beta \geq 1$, $\mathrm{A}\left(\mathrm{SNR},\lambda\right)$ is negative if and only if $\beta \leq 2$ (and positive if and only if $\beta \geq 2$), which is equivalent to the condition of Equation (3.24). Equation (3.25) is obtained using $\mathrm{SNR} = \frac{3}{2}\lambda$ in Equation (3.23). Note that SR_∞ is always positive. Since SR_∞ is an increasing function of λ if $\lambda < \frac{2}{3}\mathrm{SNR}$ and a decreasing function after this point, the maximum value of this function is given by Equation (3.25). $\square$

3.3.3 Impact of partial information on the optimal strategy

In order to measure the impact of the investor's inability to observe the trend on the optimal strategy performance, we introduce the partial information factor. This indicator represents the ratio between the asymptotic Sharpe ratio of the optimal strategy with partial information and the asymptotic Sharpe ratio of the optimal strategy with complete information:

$$\mathrm{PIF} = \frac{\mathrm{SR}_\infty}{\mathrm{SR}^{\mathrm{o}}_\infty}, \tag{3.26}$$

where SR_∞ is the asymptotic Sharpe ratio of the optimal strategy with partial information, and $\mathrm{SR}^{\mathrm{o}}_\infty$ is the asymptotic Sharpe ratio of optimal strategy with full information. The following theorem gives the analytic form of this indicator.

Theorem 3.3.4. *The partial information factor is given by:*

$$PIF = \left(\frac{\lambda}{SNR}\right)^{3/2}\frac{\left(\sqrt{1+\frac{2SNR}{\lambda}}-1\right)^2}{\sqrt{2}}, \tag{3.27}$$

where SNR is the signal-to-noise-ratio (see Equation (3.14)).
If $SNR < \frac{3}{2}\lambda$ (respectively, $SNR > \frac{3}{2}\lambda$):

1. *For a fixed parameter value SNR, this indicator is a decreasing function (respectively, an increasing function) of λ.*

2. *For a fixed parameter value λ, this indicator is an increasing function (respectively, a decreasing function) of SNR.*

Moreover:

$$PIF \leq \frac{2}{3^{3/2}}, \tag{3.28}$$

and this bound is attained for $\lambda = \frac{2}{3}SNR$.

Proof. This expression of the partial information factor is a consequence of Equations (3.13) and (3.23). Moreover:

$$\frac{\partial \mathrm{PIF}}{\partial \mathrm{SNR}} = \frac{\sqrt{\lambda}\left(\sqrt{\frac{2\mathrm{SNR}}{\lambda}+1}-1\right)\left(3\lambda\left(\sqrt{\frac{2\mathrm{SNR}}{\lambda}+1}-1\right)-2\mathrm{SNR}\right)}{2\mathrm{SNR}^{5/2}\sqrt{\frac{4\mathrm{SNR}}{\lambda}+2}}.$$

This expression is positive if and only if $\mathrm{SNR} \leq \frac{3}{2}\lambda$. The dependency on the mean reversion speed λ comes from Corollary 3.3.3. □

Remark 3.3.5. *Equation (3.28) shows that in the best configuration (with* $\lambda = \frac{2}{3}SNR$*), the asymptotic Sharpe ratio of the optimal strategy with partial information is approximatively equal to* 38.49% *of the asymptotic Sharpe ratio of the optimal strategy with complete information.*
Moreover, the intuition tells us that a high signal-to-noise ratio and a small trend mean reversion speed λ *involves a small impact of partial information on the optimal strategy performance (and then a high PIF). This is true if and only if* $SNR \leq \frac{3}{2}\lambda$.

3.4 Simulations

In this section, numerical examples are computed in order to illustrate the analytical results of the previous sections. The figure 3.1 represents the asymptotic Sharpe ratio of the optimal strategy with full information as a function of the signal-to-noise ratio. If the signal-to-noise ratio is inferior to 1, which corresponds to a trend standard deviation inferior to the volatility of the risky asset, the asymptotic Sharpe ratio of the optimal strategy with complete information is inferior to 0.5.
Now, suppose that $\lambda \in [1, 252]$ and that the trend is an unobservable process. The figure 3.2 represents the asymptotic Sharpe ratio of the optimal strategy with partial information as a function of the trend mean reversion speed λ and of the signal-to-noise ratio. Since $\lambda \in [1, 252]$ and $\mathrm{SNR} < 1$, Equation (3.24) is satisfied and this Sharpe ratio is an increasing function of SNR and a decreasing function of λ. Moreover, the maximal value is inferior to 0.2. We also observe that, even with a high signal-to-noise ratio, a high mean reversion parameter λ leads to a small Sharpe ratio.
The figure 3.3 represents the partial information factor, which corresponds to the ratio between the asymptotic Sharpe ratios of the optimal strategy with partial and full information (see Equation (3.26)). Using Equation

(3.28), this indicator is bound by $\frac{2}{3^{3/2}}$. Since SNR $< \frac{3}{2}\lambda$, this indicator is a decreasing function of λ and an increasing function of SNR. Even with a high signal-to-noise ratio, a high mean reversion parameter λ leads to a negligible performance of the optimal strategy with partial information compared to the case with full information.

The figures 3.4 and 3.5 represents the asymptotic Sharpe ratio of the optimal strategy with partial information and the partial information factor as functions of the signal-to-noise ratio and of λ with $\lambda \in [0, 2]$. Theses figures illustrate that, if SNR $> \frac{3}{2}\lambda$, these quantities are increasing functions of the trend mean reversion speed λ (and the partial information factor is also a decreasing function of the signal-to-noise ratio).

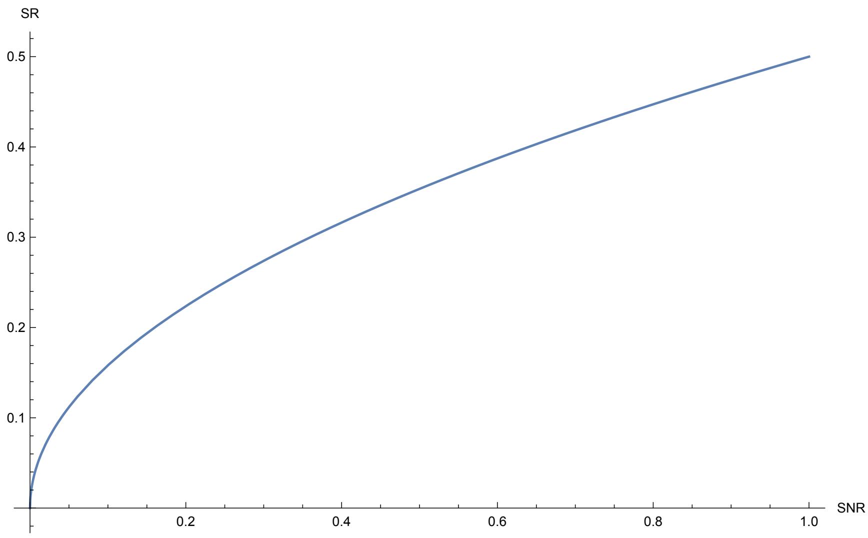

Figure 3.1: Asymptotic Sharpe ratio of the optimal strategy with complete information as a function of the signal-to-noise ratio.

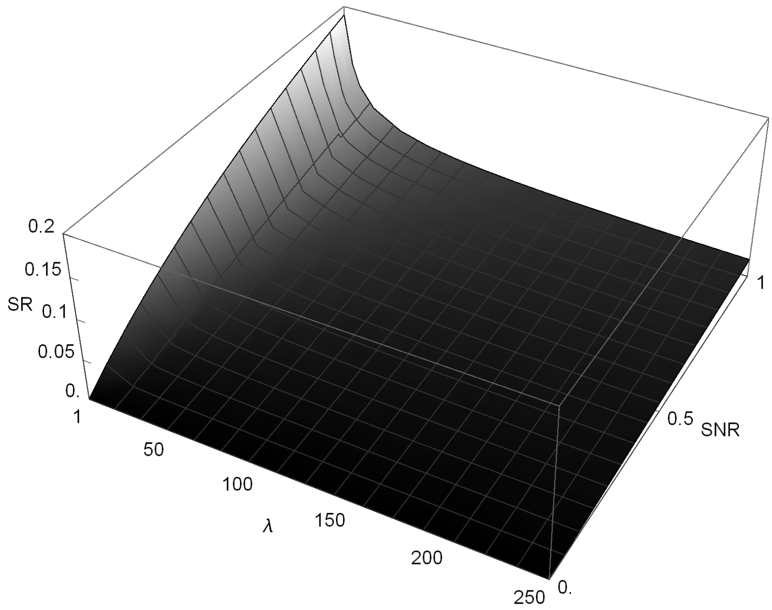

Figure 3.2: Asymptotic Sharpe ratio of the optimal strategy with partial information as a function of the trend mean reversion speed λ and of the signal-to-noise ratio.

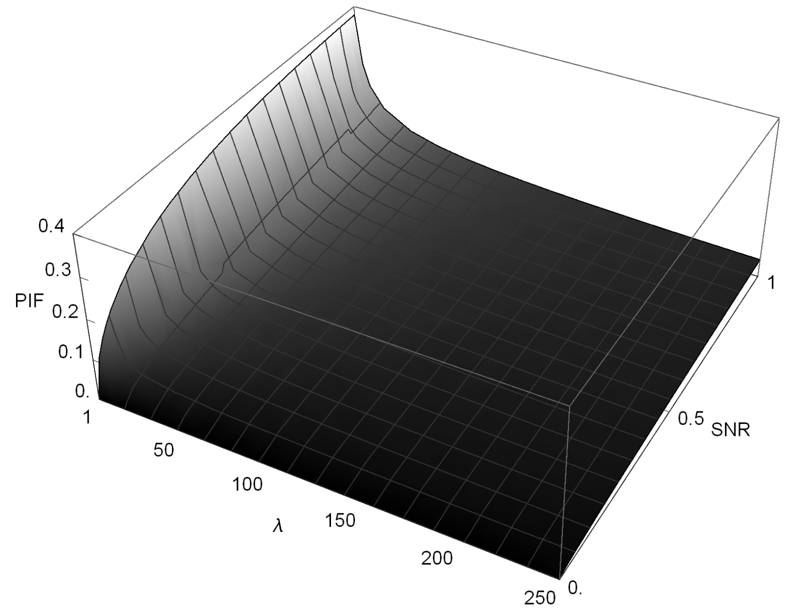

Figure 3.3: Partial information factor as a function of the trend mean reversion speed λ and of the signal-to-noise ratio.

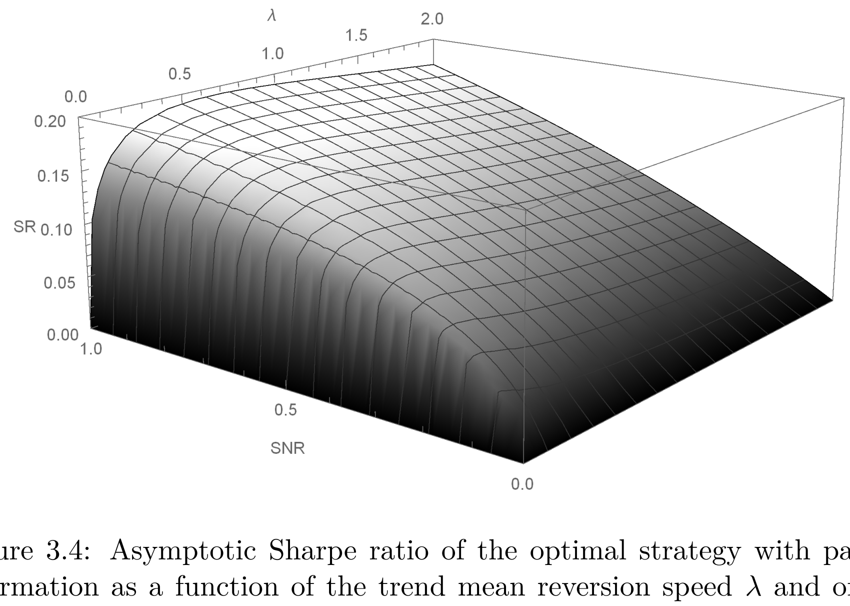

Figure 3.4: Asymptotic Sharpe ratio of the optimal strategy with partial information as a function of the trend mean reversion speed λ and of the signal-to-noise ratio with $\lambda \in [0, 2]$.

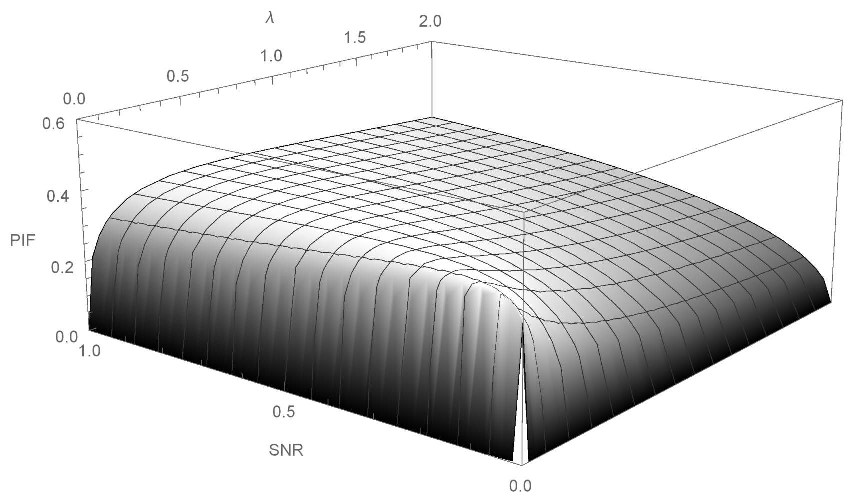

Figure 3.5: Partial information factor as a function of the trend mean reversion speed λ and of the signal-to-noise ratio with $\lambda \in [0, 2]$.

3.5 Conclusion

The present work quantifies the loss of performance in the optimal trading strategy due to partial information with a model based on an unobserved mean-reverting diffusion.
If the trend is observable, we show that the asymptotic Sharpe ratio of the optimal strategy is only an increasing function of the signal-to-noise ratio.
Under partial information, this asymptotic Sharpe ratio becomes a function of the signal-to-noise ratio and of the trend mean reversion speed. Even if the asymptotic Sharpe ratio is also an increasing function of the signal-to-noise ratio, we find that the dependency on the trend mean reversion speed is not monotonic. Indeed, this is an unimodal (increasing then decreasing) function of the trend mean reversion speed.
We also show that the ratio between the asymptotic Sharpe ratio of the optimal strategy with partial information and the asymptotic Sharpe ratio of the optimal strategy with complete information is bounded by a threshold equal to $\frac{2}{3^{3/2}}$. Given this result, we surely conclude that the impact of partial information on the optimal strategy is not negligible.
Moreover, the simulations show that even with a high signal-to-noise ratio, a high trend mean reversion speed leads to a negligible performance of the optimal strategy under partial information compared to the performance of the optimal strategy with complete information.

Chapter 4

Robustness of mathematical models and technical analysis strategies

The aim of this chapter is to compare the performances of the optimal strategy under parameters mis-specification and of a technical analysis trading strategy. The setting we consider is that of a stochastic asset price model where the trend follows an unobservable Ornstein-Uhlenbeck process. For both strategies, we provide the asymptotic expectation of the logarithmic return as functions of the model parameters. Finally, numerical examples find that an investment strategy using the cross moving averages rule is more robust than the optimal strategy under parameters misspecification.

Introduction

There exist three principal approaches for investments in financial markets (see Blanchet-Scalliet *et al.* (2007)). The first one is based on fundamental economic principles (see Tideman (1972) for details). The second one is called the technical analysis approach and uses the historical prices and volumes (see Taylor & Allen (1992),Brown & Jennings (1989) and Edwards *et al.* (2007) for details). The third one is the use of mathematical models and was introduced by Merton in 1969 (see Merton (1990) for details). He assumed that the risky asset follows a geometric Brownian motion and derived the optimal investment rules for an investor maximizing his expected utility function. Several generalisations of this problem are possible but all these models are confronted to the calibration problem. In the first chapter, we assess the feasibility of forecasting trends modeled by an unobserved mean-reverting diffusion. They show that, due to a weak signal-to-noise ratio, a bad calibration is very likely. Using this framework, Zhu and Zhou (see Zhu & Zhou (2009)) analyse the performance of a technical analysis

strategy based on a geometric moving average rule. In Blanchet-Scalliet *et al.* (2007), the authors assume that the drift is an unobservable constant piecewise process jumping at an unknown time. They provide the performance of the optimal trading strategy under parameters mis-specification and compare this strategy to a technical analysis investment based on a simple moving average rule with Monte Carlo simulations.
In this chapter, we consider a stochastic asset price model where the trend is an unobservable Ornstein Uhlenbeck process. The purpose of this work is to characterize and to compare the performances of the optimal strategy under parameters mis-specification and of a cross moving average strategy.
The chapter is organized as follows: the first section presents the model, recalls some results from filtering theory and rewrites the Kalman filter estimator as a corrected exponential average.
In the second section, the optimal trading strategy under parameters mis-specification is investigated. For this portfolio, the stochastic differential equation of the logarithmic return is found. Using this result, we provide, in closed form, the asymptotic expectation of the logarithmic return as a function of the signal-to-noise-ratio and of the trend mean reversion speed. We close this section by giving conditions on the model and the strategy parameters that guarantee a positive asymptotic expected logarithmic return and the existence of an optimal duration.
In the third section, we consider a cross moving average strategy. For this portfolio, we also provide the stochastic differential equation of the logarithmic return. We close this section by giving, in closed form, the asymptotic expectation of the logarithmic return as a function of the model parameters.
In the fourth section, numerical examples are performed. First, the best durations of the Kalman filter and of the optimal strategy under parameters mis-specification are illustrated over several trend regimes. We then compare the performances of a cross moving average strategy and of a classical optimal strategy used in the industry (with a duration $\tau = 1$ year) over several theoretical regimes. We also compare these performances under Heston's stochastic volatility model using Monte Carlo simulations. These examples show that the technical analysis approach is more robust than the optimal strategy under parameters mi-specification. We close this study by confirming this conclusion with empirical tests based on real data.

4.1 Setup

This section begins by presenting the model, which corresponds to an unobserved mean-reverting diffusion. After that, we reformulate this model in a completely observable environment (see Liptser & Shiriaev (1977) for details). This setting introduces the conditional expectation of the trend, knowing the past observations. Then, we recall the asymptotic continuous time limit of the Kalman filter and we rewrite this estimator as a corrected exponential average.

4.1.1 The model

Consider a financial market living on a stochastic basis $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$, where $\mathbf{F} = \{\mathcal{F}_t, t \geqslant 0\}$ is the natural filtration associated to a two-dimensional (uncorrelated) Wiener process (W^S, W^μ), and $\mathbb{P}$ is the objective probability measure. The dynamics of the risky asset S is given by

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_S dW_t^S, \tag{4.1}$$

$$d\mu_t = -\lambda \mu_t dt + \sigma_\mu dW_t^\mu, \tag{4.2}$$

with $\mu_0 = 0$. We also assume that $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$. The parameter λ is called the trend mean reversion speed. Indeed, λ can be seen as the "force" that pulls the trend back to zero. Denote by $\mathbf{F}^S = \left\{\mathcal{F}_t^S\right\}$ be the natural filtration associated to the price process S. An important point is that only $\mathbf{F}^S$-adapted processes are observable, which implies that agents in this market do not observe the trend μ.

4.1.2 The observable framework

As stated above, the agents can only observe the stock price process S. Since, the trend μ is not F^S-measurable, the agents do not observe it directly. Indeed, the model (4.1)-(4.2) corresponds to a system with partial information. The following proposition gives a representation of the model (4.1)-(4.2) in an observable framework (see Liptser & Shiriaev (1977) for details or Appendix E for a proof).

Proposition 23. *The dynamics of the risky asset S is also given by*

$$\frac{dS_t}{S_t} = E\left[\mu_t | \mathcal{F}_t^S\right] dt + \sigma_S dN_t, \tag{4.3}$$

where N is a $\left(\mathbb{P}, \mathbf{F}^S\right)$ Wiener process.

Remark 4.1.1. *In the filtering theory (see Liptser & Shiriaev (1977) for details), the process N is called the innovation process. To understand this name, note that:*

$$dN_t = \frac{1}{\sigma_S}\left(\frac{dS_t}{S_t} - E\left[\mu_t | \mathcal{F}_t^S\right] dt\right).$$

Then, dN_t represents the difference between the current observation and what we expect knowing the past observations.

4.1.3 Optimal trend estimator

The system (4.1)-(4.2) corresponds to a Linear Gaussian Space State model (see Brockwell & Davis (2002a) for details). In this case, the Kalman filter gives the optimal estimator, which corresponds to the conditional expectation $E\left[\mu_t|\mathcal{F}_t^S\right]$. Since $(\lambda, \sigma_\mu, \sigma_S) \in \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}_+^*$, the model (4.1)-(4.2) is a controllable and observable time invariant system. In this case, it is well known that the estimation error variance converges to an unique constant value (see Kalman *et al.* (1962) for details). This corresponds to the steady-state Kalman filter. The following proposition (see the first chapter for a proof) gives a first continuous representation of the steady-state Kalman filter:

Proposition 24. *The steady-state Kalman filter has a continuous time limit depending on the asset returns:*

$$d\widehat{\mu}_t = -\lambda\beta\widehat{\mu}_t dt + \lambda\left(\beta - 1\right)\frac{dS_t}{S_t}, \tag{4.4}$$

where

$$\beta = \left(1 + \frac{\sigma_\mu^2}{\lambda^2\sigma_S^2}\right)^{\frac{1}{2}}. \tag{4.5}$$

The steady-state Kalman filter can also be re-written as a corrected exponential average:

Proposition 25.

$$\widehat{\mu}_t = m^*\tilde{\mu}_t^*, \tag{4.6}$$

where $m^* = \frac{\beta-1}{\beta}$ *and* $\tilde{\mu}^*$ *is the exponential average given by:*

$$d\tilde{\mu}_t^* = -\frac{1}{\tau^*}\tilde{\mu}_t^* dt + \frac{1}{\tau^*}\frac{dS_t}{S_t}, \tag{4.7}$$

with an average duration $\tau^* = \frac{1}{\lambda\beta}$.

4.2 Optimal strategy under parameters mis-specification

In this section, we consider the optimal trading strategy under parameters mis-specification. For this portfolio, we first give the stochastic differential equation of the logarithmic return and we provide, in closed form, the asymptotic expectation of the logarithmic return.

4.2.1 Context

Consider the financial market defined in the first section with a risk free rate and without transaction costs. Let P be a self financing portfolio given by:

$$\begin{aligned} \frac{dP_t}{P_t} &= \omega_t \frac{dS_t}{S_t}, \\ P_0 &= x, \end{aligned}$$

where ω_t is the fraction of wealth invested in the risky asset (also named the control variable). The agent aims to maximize his expected logarithmic utility on an admissible domain $\mathcal{A}$ for the allocation process. In this section, we assume that the agent is not able to observe the trend μ. Formally, $\mathcal{A}$ represents all the $\mathbf{F}^S$-progressive and measurable processes and the problem is:

$$\omega^* = \arg \sup_{\omega \in \mathcal{A}} \mathbb{E} \left[\ln \left(P_t \right) | P_0 = x \right].$$

The solution of this problem is well known and easy to compute (see Lakner (1998) for example). Indeed, it has the following form:

$$\omega_t^* = \frac{E\left[\mu_t | \mathcal{F}_t^S\right]}{\sigma_S^2}.$$

In practice, the parameters are unknown and must be estimated. In the first chapter, we assess the feasibility of forecasting trends modeled by an unobserved mean-reverting diffusion. They show that, due to a weak signal-to-noise ratio, a bad calibration is very likely. Using Proposition 25, the steady state Kalman filter is a corrected exponential moving average of past returns. Therefore, a mis-specification on the parameters (λ, σ_μ) is equivalent to a mis-specification on the factor $\frac{\beta-1}{\beta}$ and on the duration τ^*.
Suppose that an agent thinks that the optimal duration is τ and considers:

$$\begin{aligned} d\tilde{\mu}_t &= -\frac{1}{\tau}\tilde{\mu}_t dt + \frac{1}{\tau}\frac{dS_t}{S_t}, && (4.8) \\ \tilde{\mu}_0 &= 0. && (4.9) \end{aligned}$$

Using this estimator, the agent will invest following:

$$\begin{aligned} \frac{dP_t}{P_t} &= m\frac{\tilde{\mu}_t}{\sigma_S^2}\frac{dS_t}{S_t}, && (4.10) \\ P_0 &= x, && (4.11) \end{aligned}$$

where $m > 0$. The following lemma gives the law of this filter $\tilde{\mu}$:

Lemma 4.2.1. *The exponential moving average of Equation (4.8) is given by:*

$$\tilde{\mu}_t = \frac{e^{-\frac{t}{\tau}}}{\tau} \left(\int_0^t e^{\frac{s}{\tau}} \mu_s ds + \sigma_S \int_0^t e^{\frac{s}{\tau}} dW_s^S \right). \tag{4.12}$$

Moreover, this filter is a centered Gaussian process, whose variance is:

$$\mathbb{V}\left[\tilde{\mu}_t\right] = \frac{\sigma_S^2}{2\tau} \left(1 - e^{\frac{-2t}{\tau}}\right) + \frac{\sigma_\mu^2}{\tau^2 \lambda \left(\frac{1}{\tau} - \lambda\right)} \left(\tau \frac{e^{\frac{-2t}{\tau}} - 1}{2} \right.$$
$$\left. + \frac{1 - e^{-t\left(\lambda + \frac{1}{\tau}\right)}}{\frac{1}{\tau} + \lambda} + \frac{2e^{-t\left(\lambda + \frac{3}{\tau}\right)} - e^{-2t\left(\lambda + \frac{1}{\tau}\right)} - e^{\frac{-4t}{\tau}}}{\frac{1}{\tau} - \lambda} \right).$$

Proof. Applying Itô's lemma to the function $f(\tilde{\mu}_t, t) = \tilde{\mu}_t e^{\frac{t}{\tau}}$ and using Equation (4.1), it follows that:

$$df(\tilde{\mu}_t, t) = \frac{e^{\frac{t}{\tau}}}{\tau} \left(\mu_t dt + \sigma_S dW_t^S \right).$$

The integral of this stochastic differential equation from 0 to t gives Equation (4.12). Therefore, $\tilde{\mu}$ is a Gaussian process. Its mean is null (because $\mu_0 = 0$). Since μ and W^S are supposed to be independent, the variance of the process $\tilde{\mu}$ is equal to the sum of $\mathbb{V}\left[\frac{e^{-\frac{t}{\tau}}}{\tau} \int_0^t e^{\frac{s}{\tau}} \mu_s ds\right]$ and $\mathbb{V}\left[\frac{e^{-\frac{t}{\tau}}}{\tau} \sigma_S \int_0^t e^{\frac{s}{\tau}} dW_s^S\right]$. The first term is computed using:

$$\mathbb{V}\left[\int_0^t e^{\frac{s}{\tau}} \mu_s ds\right] = \int_0^t \int_0^t e^{\frac{s_1 + s_2}{\tau}} \mathbb{E}\left[\mu_{s_1} \mu_{s_2}\right] ds_1 ds_2.$$

Since μ is a centered Ornstein Uhlenbeck, for all $s, t \geq 0$, we have:

$$\mathbb{E}\left[\mu_s \mu_t\right] = \mathbb{C}\mathrm{ov}\left[\mu_s, \mu_t\right] = \frac{\sigma_\mu^2}{2\lambda} e^{-\lambda(s+t)} \left(e^{2\lambda s \wedge t} - 1\right).$$

Finally, the second term is computed using:

$$\mathbb{V}\left[\int_0^t e^{ks} dW_s^S\right] = \frac{1}{2k} \left(e^{2kt} - 1\right),$$

with $k > 0$. □

4.2.2 Portfolio dynamic

The following proposition gives the stochastic differential equation of the mis-specified optimal portfolio:

Proposition 26. *Equation (4.10) leads to:*

$$d\ln(P_t) = \frac{m\tau}{2\sigma_S^2} d\tilde{\mu}_t^2 + m\left(\frac{\tilde{\mu}_t^2}{\sigma_S^2}\left(1 - \frac{m}{2}\right) - \frac{1}{2\tau}\right) dt. \tag{4.13}$$

Proof. Equation (4.10) is equivalent to (by Itô's lemma):

$$d\ln(P_t) = \frac{m\tilde{\mu}_t}{\sigma_S^2}\frac{dS_t}{S_t} - \frac{m^2\tilde{\mu}_t^2}{2\sigma_S^2} dt.$$

Using Equation (4.6),

$$d\ln(P_t) = \frac{m\tau}{\sigma_S^2}\tilde{\mu}_t d\tilde{\mu}_t + \frac{m\tilde{\mu}_t^2}{\sigma_S^2} - \frac{1}{2}\frac{m^2\tilde{\mu}_t^2}{\sigma_S^2} dt,$$

Itô's lemma on Equation (4.6) gives:

$$d\tilde{\mu}_t^2 = 2\tilde{\mu}_t d\tilde{\mu}_t + \frac{\sigma_S^2}{\tau^2} dt.$$

Using this equation, the dynamic of the logarithmic return follows. □

Remark 4.2.2. *Proposition 26 shows that the returns of the optimal strategy can be broken down into two terms. The first one represents an option on the square of the realized returns (called Option profile). The second term is called the Trading Impact. These terms are introduced and discussed in Bruder & Gaussel (2011) for this strategy without considering a specific diffusion for the risky asset.*

4.2.3 Expected logarithmic return

The following theorem gives the asymptotic expected logarithmic return of the mis-specified optimal strategy.

Theorem 4.2.3. *Consider the portfolio given by Equation (4.10). In this case:*

$$\lim_{T\to\infty} \frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T} = m\frac{\tau\left(\beta^2 - 1\right)(2 - m) - m\left(\tau + \frac{1}{\lambda}\right)}{4\tau\left(\tau + \frac{1}{\lambda}\right)}, \tag{4.14}$$

where β is given by Equation (4.5).

Proof. Using Proposition 26, it follows that:

$$\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right] = \frac{m\tau}{2\sigma_S^2}\mathbb{E}\left(\tilde{\mu}_T\right)^2 + m\int_0^T \left(\frac{\mathbb{E}\left(\tilde{\mu}_t\right)^2(2 - m)}{2\sigma_S^2} - \frac{1}{2\tau}\right) dt.$$

Moreover, $\mathbb{E}\left[\left(\tilde{\mu}_t\right)^2\right]$ is given by Lemma 4.2.1. Then, integrating the expression from 0 to T and tending T to ∞, the result follows. □

The following result is a corollary of the previous theorem. It represents the asymptotic expected logarithmic return as a function of the signal-to-noise-ratio and of the trend mean reversion speed λ.

Corollary 4.2.4. *Consider the portfolio given by Equation (4.10). In this case:*

$$\lim_{T\to\infty}\frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T}=m\frac{2\tau\,(2-m)\,SNR-m\,(\lambda\tau+1)}{4\tau\,(\lambda\tau+1)},\tag{4.15}$$

where SNR is the signal-to-noise-ratio:

$$SNR=\frac{\sigma_\mu^2}{2\lambda\sigma_S^2}.\tag{4.16}$$

Moreover:

1. *If $m<2$, for a fixed parameter value λ, this asymptotic expected logarithmic return is an increasing function of SNR.*
2. *For a fixed parameter value SNR, it is a decreasing function of λ.*

Proof. Since $\beta=\sqrt{1+\frac{2\mathrm{SNR}}{\lambda}}$, the use of this expression in Equation (4.14) gives the result. □

Remark 4.2.5. *Assume that the agent makes a good calibration and uses $m^*=\frac{\beta-1}{\beta}$ and $\tau^*=\frac{1}{\lambda\beta}$. In this case, we obtain the result of the second chapter:*

$$\lim_{T\to\infty}\frac{\mathbb{E}\left[\ln\left(\frac{P_T}{P_0}\right)\right]}{T}=\frac{1}{2}\left(SNR+\lambda-\sqrt{\lambda\,(\lambda+2SNR)}\right),\tag{4.17}$$

where SNR is defined in Equation (4.16).

The following proposition gives conditions on the trend parameters and on the duration τ that guarantee a positive asymptotic expected logarithmic return and the existence of an optimal duration.

Proposition 27. *Consider the portfolio given by Equation (4.10) and suppose that $m<2$. In this case, the asymptotic expected logarithmic return is positive if and only if:*

1. $\frac{SNR}{\lambda}>\frac{2m}{2-m}$.
2. $\tau>\tau_{min}$, *where:*

$$\tau_{min}=\frac{m}{2\,(2-m)\,SNR-\lambda m}.\tag{4.18}$$

Moreover, there exists an optimal duration $\tau_{min} < \tau_{opt} < \infty$ *if and only if* $\frac{SNR}{\lambda} > \frac{2m}{2-m}$ *and:*

$$\tau_{opt} = \frac{m + \sqrt{(2-m)\, 2m \frac{SNR}{\lambda}}}{2\,(2-m)\, SNR - \lambda m}. \tag{4.19}$$

Proof. Using Equation (4.15), the first part of the proposition follows. Since the asymptotic expected logarithmic return of the mis-specified strategy is positive after τ_{min} and tends to zero if τ tends to the infinity, there exists an optimal duration τ_{opt}. This point is computed with setting to zero the derivative of Equation (4.15) with respect to the parameter τ. $\square$

4.3 cross moving average strategy

In this section, we consider a cross moving average strategy based on geometric moving averages. For this portfolio, we first give the stochastic differential equation of the logarithmic return and we provide, in closed form, the asymptotic expectation of the logarithmic return.

4.3.1 Context

Consider the financial market defined in the first section with a risk free rate and without transaction costs. Let $G\,(t, L)$ be the geometric moving average at time t of the stock prices on a window L:

$$G\,(t, L) = \exp\left(\frac{1}{L}\int_{t-L}^{t} \log\left(S_u\right) du\right). \tag{4.20}$$

Let Q be a self financing portfolio given by:

$$\frac{dQ_t}{Q_t} = \theta_t \frac{dS_t}{S_t}, \tag{4.21}$$

$$Q_0 = x, \tag{4.22}$$

where θ_t is the fraction of wealth invested by the agent in the risky asset:

$$\theta_t = \gamma + \alpha\, \mathbf{1}_{G(t,L_1) > G(t,L_2)}$$

with $\gamma, \alpha \in \mathbb{R}$ and $0 < L_1 < L_2 < t$. This trading strategy is a combination of a fixed strategy and a pure cross moving average strategy.

4.3.2 Portfolio dynamic

The following proposition gives the stochastic differential equation of the cross moving average portfolio.

Proposition 28. *Equation (4.21) leads to:*

$$\begin{aligned}
d\ln(Q_t) &= \Bigg(\left(\gamma+\alpha\,\mathbf{1}_{G(t,L_1)>G(t,L_2)}\right)\mu_t-\frac{\gamma^2\sigma_S^2}{2}\\
&\quad-\frac{\left(\alpha^2+2\alpha\gamma\right)\sigma_S^2}{2}\,\mathbf{1}_{G(t,L_1)>G(t,L_2)}\Bigg)\,dt\\
&\quad+\left(\gamma+\alpha\,\mathbf{1}_{G(t,L_1)>G(t,L_2)}\right)\sigma_S dW_t^S.
\end{aligned}$$

Proof. Applying Itô's lemma to the process $\ln(Q)$ and using

$$\mathbf{1}^2_{G(t,L_1)>G(t,L_2)}=\mathbf{1}_{G(t,L_1)>G(t,L_2)},$$

Proposition 28 follows. □

4.3.3 Expected logarithmic return

The following theorem gives the asymptotic expected logarithmic return of the cross moving average portfolio.

Theorem 4.3.1. *Consider the portfolio given by Equation (4.21). In this case:*

$$\begin{aligned}
\lim_{T\to\infty}\frac{\mathbb{E}\left[\ln\left(\frac{Q_T}{Q_0}\right)\right]}{T} &= -\frac{\gamma^2\sigma_S^2}{2}-\frac{\left(\alpha^2+2\alpha\gamma\right)\sigma_S^2}{2}\Phi\left(\frac{m_{(L_1,L_2,\sigma_S)}}{\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\right)\\
&+\frac{\alpha\sigma_\mu^2\left(L_2\left(1-e^{-\lambda L_1}\right)-L_1\left(1-e^{-\lambda L_2}\right)\right)}{2\lambda^3L_1L_2\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\Phi'\left(-\frac{m_{(L_1,L_2,\sigma_S)}}{\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\right),
\end{aligned}$$

where Φ is the cumulative distribution function of the standard normal variable and:

$$\begin{aligned}
m_{(L_1,L_2,\sigma_S)} &= \frac{-\sigma_S^2}{4}\left(L_2-L_1\right),\\
s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)} &= \left(\frac{\sigma_\mu^2}{\lambda^2}+\sigma_S^2\right)\frac{\left(L_2-L_1\right)^2}{3L_2}-\frac{\sigma_\mu^2}{\lambda^4}\left(\frac{1}{L_1}-\frac{1}{L_2}\right)\\
&\quad+\frac{\sigma_\mu^2}{\lambda^5}\Bigg[\frac{1}{L_1^2}\left(1-e^{-\lambda L_1}\right)+\frac{1}{L_2^2}\left(1-e^{-\lambda L_2}\right)\\
&\quad-\frac{1}{L_1L_2}\left(1-e^{-\lambda L_1}\right)\left(1-e^{-\lambda L_2}\right)\\
&\quad-\frac{1}{L_1L_2}\left(e^{-\lambda(L_2-L_1)}-e^{-\lambda(L_2+L_1)}\right)\Bigg].
\end{aligned}$$

Proof. Since the processes μ and W^S are centered, Proposition 28 implies that:

$$\begin{aligned}\mathbb{E}\left[\ln\left(\frac{Q_T}{Q_0}\right)\right] &= \frac{-\gamma^2\sigma_S^2}{2}\left(T-L_2\right)\\ &\quad+\alpha\int_{L_2}^{T}\mathbb{E}\left[\mu_t\,\mathbf{1}_{G(t,L_1)>G(t,L_2)}\right]dt\\ &\quad-\frac{\left(\alpha^2+2\alpha\gamma\right)\sigma_S^2}{2}\int_{L_2}^{T}\mathbb{E}\left[\mathbf{1}_{G(t,L_1)>G(t,L_2)}\right]dt,\end{aligned}$$

where $T > L_2$. Let $t > L_2$ and consider the following process:

$$X_t = m_1(t) - m_2(t), \tag{4.23}$$

where $\forall i \in \{1, 2\}$:

$$m_i(t) = \frac{1}{L_i}\int_{t-L_i}^{t}\log\left(S_u\right)du.$$

Then X is a Gaussian process. Based on Lemma 2 in Zhu & Zhou (2009), $\forall t > L_2$:

$$\{G(t,L_1) > G(t,L_2)\} \Leftrightarrow \{X_t > 0\}, \tag{4.24}$$

$$\mathbb{E}\left[\mathbf{1}_{G(t,L_1)>G(t,L_2)}\right] = \Phi\left(\frac{\mathbb{E}[X_t]}{\sqrt{\mathbb{V}ar[X_t]}}\right), \tag{4.25}$$

$$\mathbb{E}\left[\mu_t\,\mathbf{1}_{G(t,L_1)>G(t,L_2)}\right] = \frac{\mathbb{C}ov[X_t,\mu_t]}{\sqrt{\mathbb{V}ar[X_t]}}\Phi'\left(-\frac{\mathbb{E}[X_t]}{\sqrt{\mathbb{V}ar[X_t]}}\right) \tag{4.26}$$

The following lemma gives the mean, the asymptotic variance of the process X and the covariance function between the processes X and μ.

Lemma 4.3.2. *Consider the process X defined in Equation (4.23). In this case, $\forall t > L_2$:*

$$\mathbb{E}[X_t] = \frac{-\sigma_S^2}{4}\left(L_2 - L_1\right), \tag{4.27}$$

$$\lim_{t\to\infty}\mathbb{V}ar[X_t] = s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}, \tag{4.28}$$

$$\mathbb{C}ov[X_t,\mu_t] = g(t,L_1) - g(t,L_2), \tag{4.29}$$

where $s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}$ is defined in Theorem 4.3.1 and

$$g(t,L) = \frac{-\sigma_\mu^2 e^{-\lambda t}}{\lambda^2 L}\left(\lambda L + \sinh\left(\lambda\left(t-L\right)\right) - \sinh\left(\lambda t\right)\right). \tag{4.30}$$

Proof of Lemma 4.3.2. Since:

$$\mathbb{E}\left[m_i\left(t\right)\right] = \frac{-\sigma_S^2}{4}\left(2t - L_i\right),$$

Equation (4.27) follows. Moreover:

$$\mathbb{C}ov\left[m_1\left(t\right), m_2\left(t\right)\right] = \frac{1}{L_1 L_2}\int_{t-L_1}^{t}\int_{t-L_2}^{t}\mathbb{C}ov\left[\ln S_u, \ln S_v\right]dudv,$$

Since

$$\mathbb{C}ov\left[\ln S_u, \ln S_v\right] = \int_0^u\int_0^v \mathbb{C}ov\left[\mu_s, \mu_t\right]dsdt + \sigma_S^2 \min\left(u, v\right),$$

and the drift μ is an Ornstein Uhlenbeck process:

$$\mathbb{C}ov\left[\mu_s, \mu_t\right] = \frac{\sigma_\mu^2 e^{-\lambda(s+t)}}{2\lambda}\left(e^{2\lambda\min(s,t)} - 1\right).$$

Then

$$\begin{aligned}\mathbb{C}ov\left[\ln S_u, \ln S_v\right] &= \left(\sigma_S^2 + \frac{\sigma_\mu^2}{\lambda^2}\right)\min\left(u, v\right)\\ &\quad + \frac{\sigma_\mu^2}{2\lambda^3}\left(2e^{-\lambda u} + 2e^{-\lambda v} - e^{-\lambda|v-u|} - e^{-\lambda(v+u)} - 1\right).\end{aligned}$$

Using

$$\mathbb{V}ar\left[X_t\right] = \mathbb{V}ar\left[m_1\left(t\right)\right] + \mathbb{V}ar\left[m_2\left(t\right)\right] - 2\mathbb{C}ov\left[m_1\left(t\right), m_2\left(t\right)\right]$$

and tending t to ∞ Equation (4.28) follows. Since the processes W^S and μ are supposed to be independent, there holds:

$$\mathbb{C}ov\left[X_t, \mu_t\right] = \mathbb{C}ov\left[m_1\left(t\right), \mu_t\right] - \mathbb{C}ov\left[m_2\left(t\right), \mu_t\right].$$

Moreover

$$\mathbb{C}ov\left[m_i\left(t\right), \mu_t\right] = \frac{1}{L_i}\int_{t-L_i}^{t}\mathbb{C}ov\left[\ln S_u, \mu_t\right]du,$$

and

$$\mathbb{C}ov\left[\ln S_u, \mu_t\right] = \int_0^u \mathbb{C}ov\left[\mu_s, \mu_t\right]ds,$$

then

$$\mathbb{C}ov\left[m_i\left(t\right), \mu_t\right] = g\left(t, L_i\right),$$

where the function g is defined in Equation (4.30). Equation (4.29) follows $\square$

The use of Lemma 4.3.2 gives:

$$
\begin{aligned}
\mathbb{E}\left[\ln\left(\frac{Q_T}{Q_0}\right)\right] &= \frac{-\gamma^2\sigma_S^2}{2}(T-L_2) \\
&+\alpha\Phi'\left(-\frac{m_{(L_1,L_2,\sigma_S)}}{\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\right)\int_{L_2}^{T}\frac{\mathbb{C}ov\left[X_t,\mu_t\right]}{\sqrt{\mathbb{V}ar\left[X_t\right]}}dt \\
&-\frac{\left(\alpha^2+2\alpha\gamma\right)\sigma_S^2}{2}(T-L_2)\,\Phi\left(\frac{m_{(L_1,L_2,\sigma_S)}}{\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}}\right).
\end{aligned}
$$

Moreover, a direct calculus shows that:

$$
\lim_{T\to\infty}\frac{\int_{L_2}^{T}\frac{\mathbb{C}ov[X_t,\mu_t]}{\sqrt{\mathbb{V}ar[X_t]}}dt}{T} = \frac{\sigma_\mu^2\left(L_2\left(1-e^{-\lambda L_1}\right)-L_1\left(1-e^{-\lambda L_2}\right)\right)}{2\lambda^3 L_1 L_2\sqrt{s_{(L_1,L_2,\lambda,\sigma_\mu,\sigma_S)}}},
$$

the result of Theorem 4.3.1 follows. □

4.3.4 Strategy with one moving average

Suppose that $L_1 = 0$ and $L_2 = L$. In this case, the fraction of wealth invested by the agent in the risky asset becomes:

$$
\theta_t^1 = \gamma + \alpha\,\mathbf{1}_{S_t > G(t,L)},
$$

where G is the geometric moving average defined in Equation (4.20) and the self financing portfolio Q^1 becomes:

$$
\begin{aligned}
\frac{dQ_t^1}{Q_t^1} &= \theta_t^1\frac{dS_t}{S_t}, && (4.31)\\
Q_0^1 &= x, && (4.32)
\end{aligned}
$$

This particular case corresponds to the allocation introduced in Zhu & Zhou (2009) when we assume that the two Brownian motions W^S and W^μ are uncorrelated and that the trend is mean reverted around 0. Given this framework, we can provide the asymptotic expected logarithmic return of this trading strategy (which has already been found in Zhu & Zhou (2009)):

Theorem 4.3.3. *Consider the portfolio given by Equation (4.31). In this case:*

$$
\begin{aligned}
\lim_{T\to\infty}\frac{\mathbb{E}\left[\ln\left(\frac{Q_T^1}{Q_0^1}\right)\right]}{T} &= -\frac{\gamma^2\sigma_S^2}{2}-\frac{\left(\alpha^2+2\alpha\gamma\right)\sigma_S^2}{2}\Phi\left(\frac{m^1_{(L,\sigma_S)}}{\sqrt{s^1_{(L,\lambda,\sigma_\mu,\sigma_S)}}}\right)\\
&+\frac{\alpha\sigma_\mu^2\frac{1-\left(1-e^{-\lambda L}\right)}{\lambda L}}{2\lambda^2\sqrt{s^1_{(L,\lambda,\sigma_\mu,\sigma_S)}}}\Phi'\left(-\frac{m^1_{(L,\sigma_S)}}{\sqrt{s^1_{(L,\lambda,\sigma_\mu,\sigma_S)}}}\right),
\end{aligned}
$$

where Φ is the cumulative distribution function of the standard normal variable and:

$$\begin{aligned}
m^1_{(L,\sigma_S)} &= m_{(0,L,\sigma_S)} \\
&= \frac{-\sigma_S^2}{4}L, \\
s^1_{(L,\lambda,\sigma_\mu,\sigma_S)} &= s_{(0,L,\lambda,\sigma_\mu,\sigma_S)} \\
&= \frac{\left(\frac{\sigma_\mu^2}{\lambda^2}+\sigma_S^2\right)L}{3} - \frac{\sigma_\mu^2}{2\lambda^3}\left(1-\frac{2\left(1-e^{-\lambda L}\left(1+\lambda L\right)\right)}{\lambda^2 L^2}\right),
\end{aligned}$$

and the functions s and m are introduced in Theorem 4.3.1.

Proof. This result is a consequence of Theorem 4.3.1. Indeed, tending L_1 to 0 and using $L_2 = L$, the result follows. □

4.4 Simulations

In this section, numerical simulations and empirical tests based on real data are performed. The aim of these tests is to compare the robustness of the optimal strategy under parameters mis-specification and of an investment using cross moving averages. First, the best durations of the Kalman filter and of the optimal strategy under parameters mis-specification are illustrated over several trend regimes. We then consider the asymptotic expected logarithmic returns of the cross moving average strategy (see Section 4.3) with $(L_1, L_2) = (5 \text{ days}, 252 \text{ days})$ and of the optimal strategy with a duration $\tau = 252$ days. Using this configuration, we study the stability of the performances of these strategies over several theoretical regimes. We also confirm our results under Heston's stochastic volatility model with Monte Carlo simulation. Finally, backtests of these two strategies on real data confirm our theoretical expectations.

4.4.1 Optimal durations

In this subsection, we consider the model (4.1)-(4.2).

Well-specified Kalman filter

In these simulations, we consider a signal-to-noise ratio inferior to 1. This assumption corresponds to a trend standard deviation inferior to the volatility of the risky asset. Using $\tau^* = \frac{1}{\lambda\beta}$ and $\beta = \sqrt{1+\frac{2\mathrm{SNR}}{\lambda}}$, The figures 4.1 and 4.2 represent the optimal Kalman filter duration τ^* as a function of the trend mean reversion speed λ and of the signal-to-noise ratio. This duration is a decreasing function of these parameters. Indeed, if the variation of the trend process is low and if the measurement noise is high compared to the trend standard deviation, the window of filtering must be long. Moreover, we observe that for a trend mean reversion speed inferior to 1 (which corresponds to a slow trend process), the duration τ^* is superior to 0.5 years and can reach 10 years. If the trend mean reversion speed is superior to 1, this duration is inferior to 1 year.

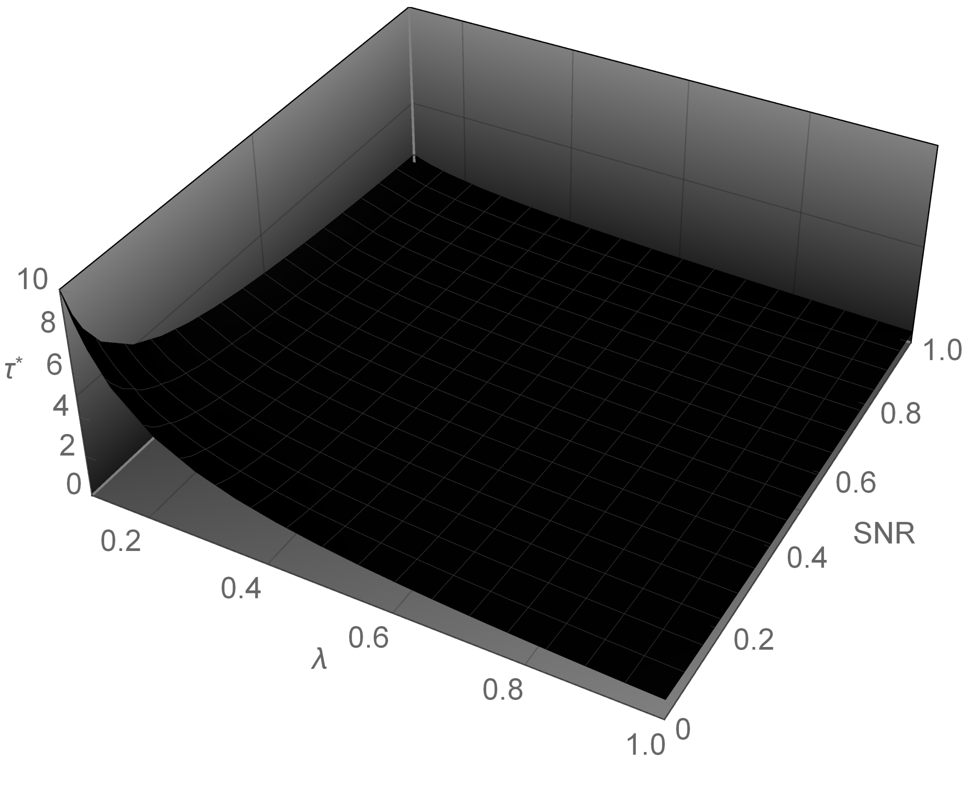

Figure 4.1: Optimal duration (years) of the Kalman filter with $\lambda \in [0.1, 1]$

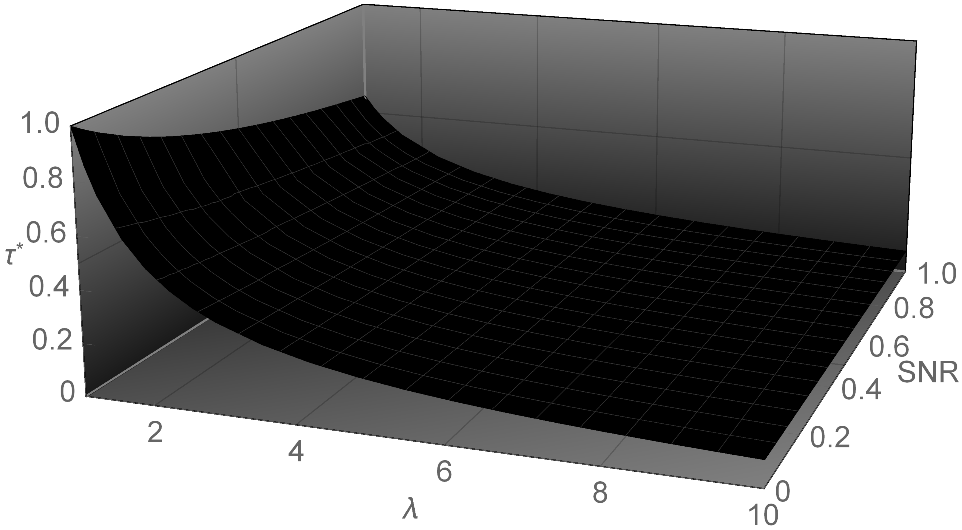

Figure 4.2: Optimal duration (years) of the Kalman filter with $\lambda \in [1, 10]$

Best filtering window for the optimal strategy under parameters mis-specification

Under parameters mis-specification, we can also define an optimal duration using the strategy introduced in Section 4.2 and Proposition 27. This duration is the one maximizing the asymptotic expected logarithmic return of the optimal strategy under parameters mis-specification. This optimal window exists if and only if $\frac{\mathrm{SNR}}{\lambda} > \frac{2m}{2-m}$. We assume that $m = 1$. Then, the condition becomes $\frac{\mathrm{SNR}}{\lambda} > 2$. The figures 4.3 and 4.4 represent this duration $\tau_{\mathrm{opt}}\,(m = 1)$ as a function of the trend mean reversion speed λ with respectively SNR$= 1$ and SNR$= 0.5$. This duration has a similar behaviour than the optimal Kalman filter duration, except when the trend mean reversion speed λ tends to $\frac{\mathrm{SNR}}{2}$. Indeed, if $\lambda = \frac{\mathrm{SNR}}{2}$, the condition $\frac{\mathrm{SNR}}{\lambda} > 2$ is not satisfied and the optimal duration becomes infinite.

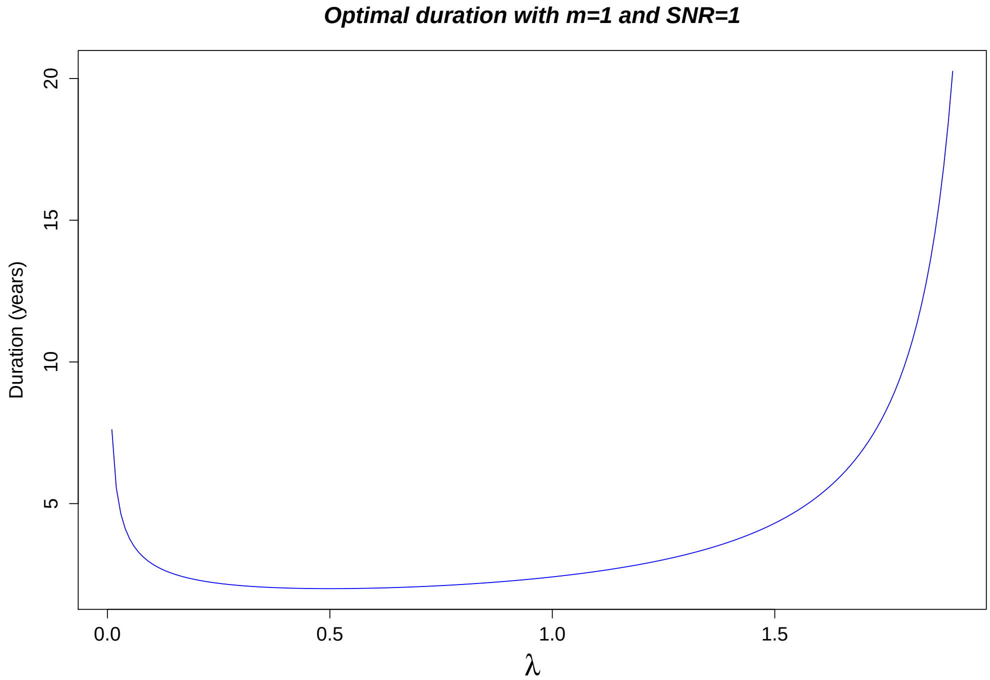

Figure 4.3: Optimal duration (years) of the mis-specified filter with $m = 1$ and SNR$= 1$

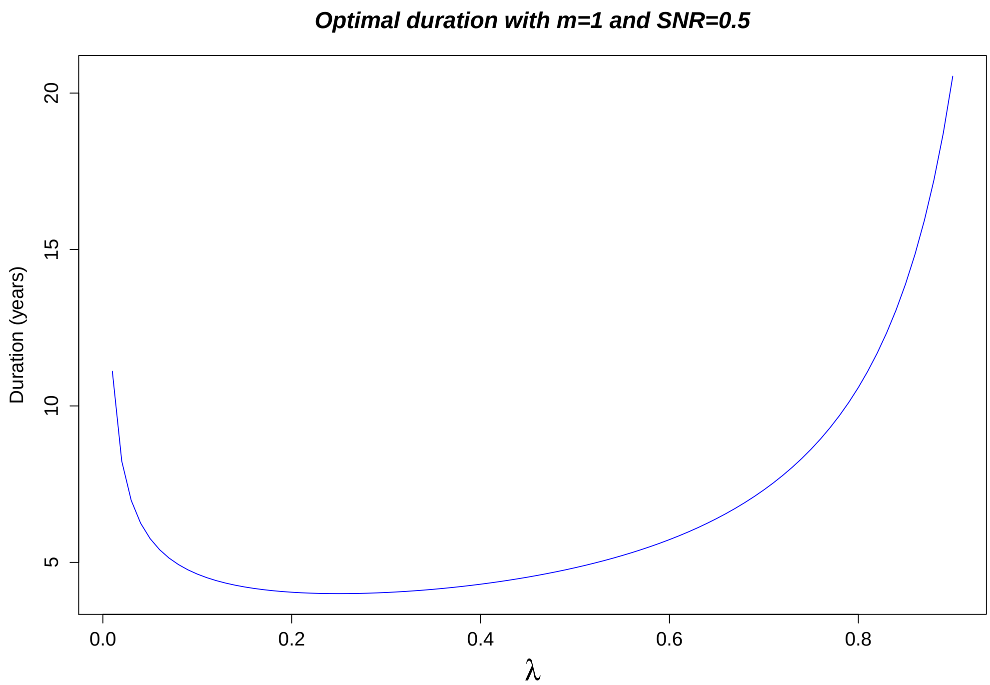

Figure 4.4: Optimal duration (years) of the mis-specified filter with $m = 1$ and SNR$= 0.5$

4.4.2 Robustness of the optimal strategy and of the cross moving average strategy

Stability of the performances over several theoretical regimes under constant spot volatility

In this subsection, we consider the model (4.1)-(4.2). Moreover, we assume that a year contains 252 days and that the risky asset volatility is equal to $\sigma_S = 30\%$. We consider two trading strategies. The first one is the optimal strategy (introduced in section 4.2) with a duration $\tau = 252$ days ($= 1$ year) and a leverage $m = 1$. The second strategy is the cross moving average strategy (introduced in section 4.3) with $(L_1, L_2) = (5 \text{ days}, 252 \text{ days})$ and the following allocation:

$$\theta_t = -1 + 2\,\mathbf{1}_{G(t,L_1)>G(t,L_2)},$$

where G is the geometric moving average defined in Equation (4.20). Then, if the short geometric average is superior (respectively inferior) to the long geometric average, we buy (respectively sell) the risky asset. In order to compare the performance stability of these two strategies, we use the asymptotic expected logarithmic returns found in Theorems 4.2.3 and 4.3.1. The figures 4.5, 4.6, 4.7 and 4.8 represent the performances of these strategies after 100 years as a function of the trend volatility σ_μ respectively with $\lambda = 1, 2, 3$ and 4. Even if the optimal strategy can provide a better performance (for example with $\lambda = 1$ and $\sigma_\mu = 90\%$), it can also provide higher losses than the cross average strategy (for example with $\lambda = 4$ and $\sigma_\mu = 10\%$). We can conclude with these tests that the theoretical performance of this cross average strategy is more robust than the theoretical performance of this optimal strategy.

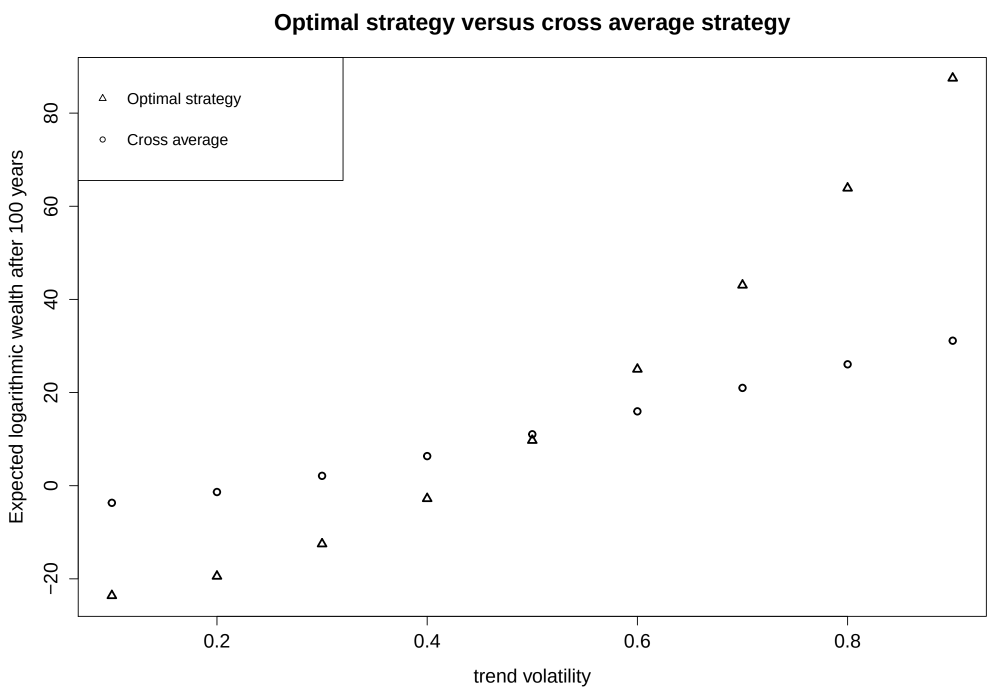

Figure 4.5: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 1$, $\sigma_S = 30\%$ and $T = 100$ years

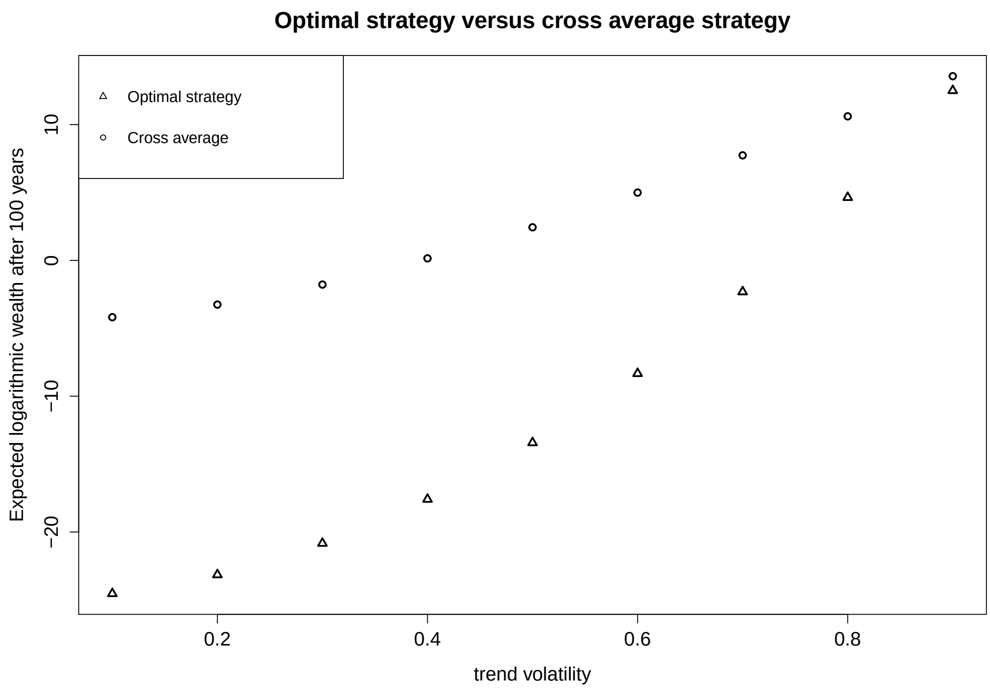

Figure 4.6: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 2$, $\sigma_S = 30\%$ and $T = 100$ years

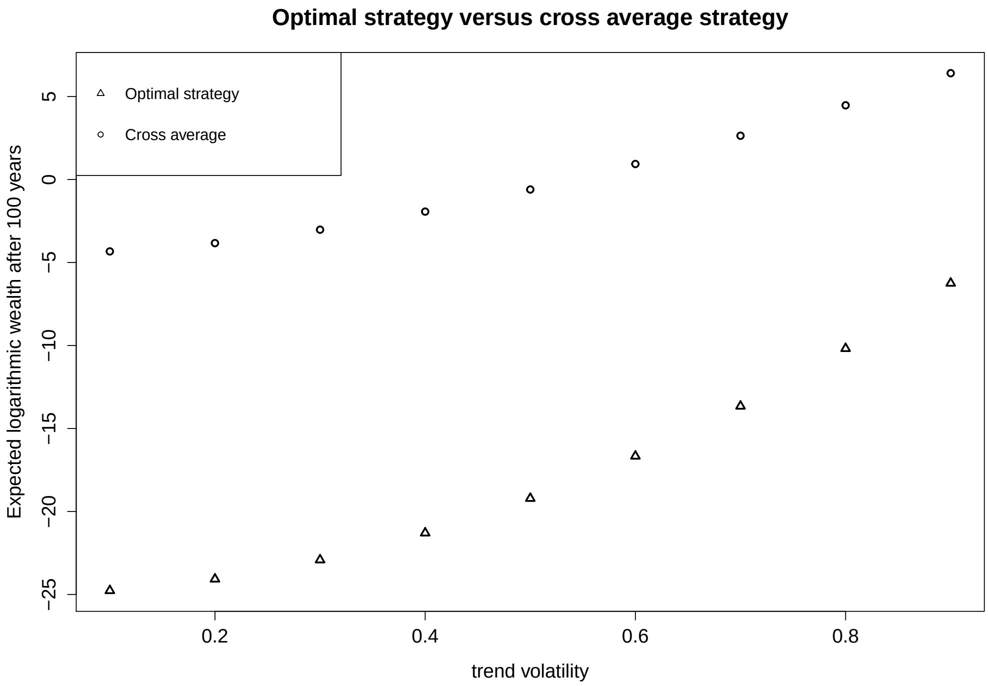

Figure 4.7: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 3$, $\sigma_S = 30\%$ and $T = 100$ years

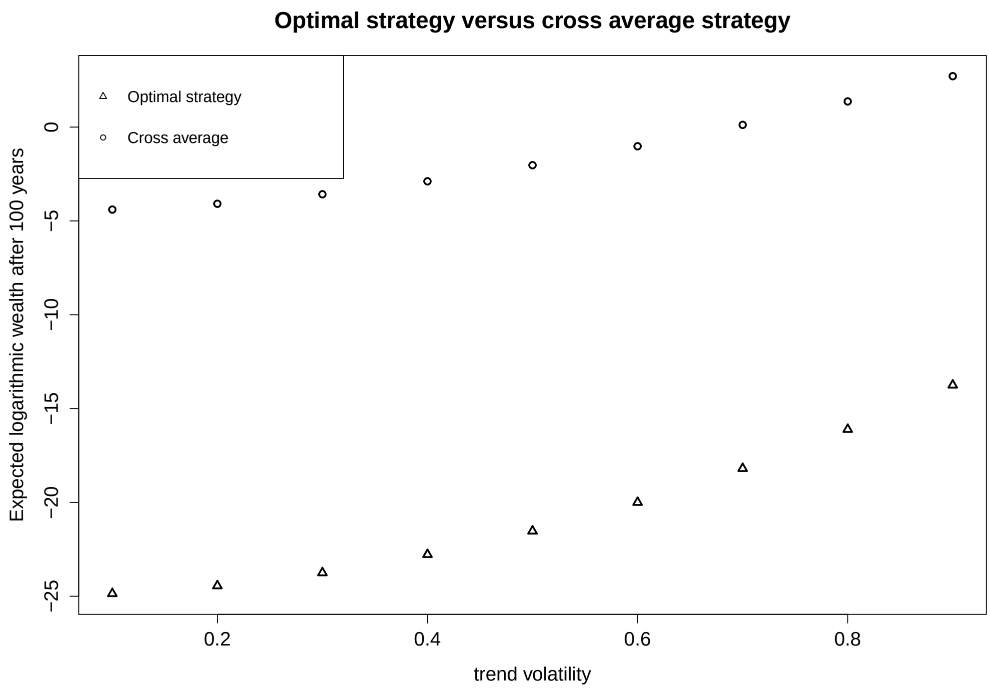

Figure 4.8: The expected logarithmic returns of the optimal strategy (with $\tau = 252$ days) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $\lambda = 4$, $\sigma_S = 30\%$ and $T = 100$ years

Stability of the performances over several theoretical regimes under Heston's stochastic volatility model

Model and optimal strategy The aim of this subsection is to check if the cross average strategy is more robust than the optimal trading strategy under Heston's stochastic volatility model (see Heston (1993) or Mikhailov & Nögel (2003) for details). To this end, consider a financial market living on a stochastic basis $(\Omega, \mathcal{G}, \mathbf{G}, \mathbb{P})$, where $\mathbf{G} = \{\mathcal{G}_t, t \geqslant 0\}$ is the natural filtration associated to a three-dimensional Wiener process (W^S, W^μ, W^V), and $\mathbb{P}$ is the objective probability measure. The dynamics of the risky asset S is given by

$$\frac{dS_t}{S_t} = \mu_t dt + \sqrt{V_t} dW_t^S, \tag{4.33}$$

$$d\mu_t = -\lambda \mu_t dt + \sigma_\mu dW_t^\mu, \tag{4.34}$$

$$dV_t = \alpha \left(V_\infty - V_t \right) dt + \epsilon \sqrt{V_t} dW_t^V \tag{4.35}$$

with $\mu_0 = 0$, $V_0 > 0$, $d\left\langle W^S, W^\mu \right\rangle_t = 0$, and $d\left\langle W^S, W^V \right\rangle_t = \rho dt$. We also assume that $(\lambda, \sigma_\mu) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$ and that $2kV_\infty > \epsilon$ (in this case, the variance V cannot reach zero and is always positive, see Cox *et al.* (1985) for details). Denote by $\mathbf{G}^S = \left\{\mathcal{G}_t^S\right\}$ be the natural filtration associated to the price process S. In this case, the process V is G^S-adapted. Now, assume that the agent aims to maximize his expected logarithmic wealth (on an admissible domain $\mathcal{A}$, which represents all the $\mathbf{G}^S$-progressive and measurable processes). In this case, his optimal portfolio is given by (see Bjork *et al.* (2010)):

$$\begin{aligned}
\frac{dP_t}{P_t} &= \frac{E\left[\mu_t | \mathcal{G}_t^S\right]}{V_t} \frac{dS_t}{S_t}, \\
P_0 &= x.
\end{aligned}$$

Let δ be a discrete time step, and denote by the subscript k the value of a process at time $t_k = k\delta$. Using the scheme that produces the smallest discretization bias for the variance process (see Lord *et al.* (2010) for details), the discrete time model is:

$$y_{k+1} = \frac{S_{k+1} - S_k}{\delta S_k} = \mu_{k+1} + u_{k+1}, \tag{4.36}$$

$$\mu_{k+1} = e^{-\lambda\delta}\mu_k + v_k, \tag{4.37}$$

$$V_{k+1} = V_k + \alpha\left(V_\infty - V_k^+\right)\delta + \epsilon\sqrt{V_k^+} z_k \tag{4.38}$$

where $x^+ = \max(0, x)$, $u_{k+1} \sim \mathcal{N}\left(0, \frac{V_k}{\delta}\right)$, $v_k \sim \mathcal{N}\left(0, \frac{\sigma_\mu^2}{2\lambda}\left(1 - e^{-2\lambda\delta}\right)\right)$ and $z_k \sim \mathcal{N}(0, \delta)$.

Monte Carlo simulations In this section, Monte Carlo simulations are used to check if the cross average strategy is more robust than the optimal trading strategy under Heston's stochastic volatility model. To this end, we consider the discrete model (4.36)-(4.37)-(4.38) and we assume that $\alpha = 4$ (quarterly mean-reversion of the variance process), that $\epsilon = 5\%$, that $V_\infty = V_0 = 0.3^2$ (which means an initial and a long horizon spot volatility equal to 30%) and that $\rho = -60\%$ (when the spot decreases, the volatility increases). Moreover, we consider an investment horizon equal to 50 years and $\delta = 1/252$ (which means that that a year contains 252 days and that each allocation is made daily). With this set-up, we consider several trend regimes, we simulate M paths of the risky asset over 50 years and we implement two strategies:

1. The discrete time version of the optimal strategy presented above. Since the process V is G^S-adapted, V_k is observable at time t_k and the

conditional expectation of the trend is tractable with the non stationary discrete time Kalman filter (see Appendix A). We assume that the agent thinks that the parameters are equal to $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$ when he uses the Kalman filter.

2. The cross moving average strategy (introduced in section 4.3) with $(L_1, L_2) = (5 \text{ days}, 252 \text{ days})$ and the following allocation:

$$\theta_k = -1 + 2\,\mathbf{1}_{G^d(k,L_1) > G^d(k,L_2)},$$

where $G^d(k, L)$ is the discrete geometric moving average computed on the last L values of S.

The figures 4.9 and 4.10 represent the estimated performances of these strategies after 50 years as a function of the trend volatility σ_μ with $M = 10000$ and respectively with $\lambda = 1$ and 2. These results confirm that the performance of the cross average strategy is less sensitive to a trend regime variation than the performance optimal trading strategy with parameters mis-specification. Moreover, The figures 4.11, 4.12, 4.13 and 4.14 represent the empirical distribution of the logarithmic return of these strategies after 50 years over $M = 10000$ paths for different configurations. These figures show that, even with a good calibration, the logarithmic return of the cross average strategy is less dispersed than the logarithmic return of the optimal strategy. Then the cross average strategy is more robust than the optimal strategy.

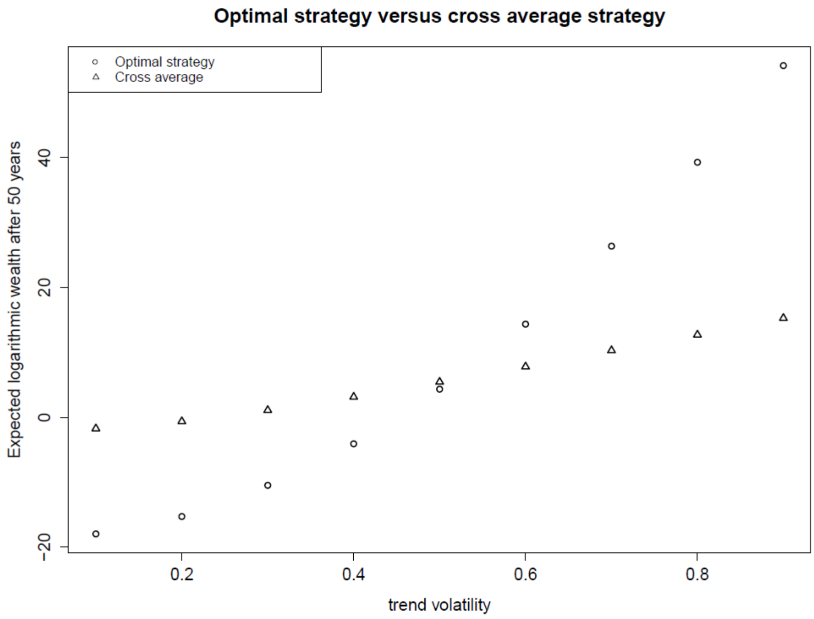

Figure 4.9: The expected logarithmic returns of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) as functions of σ_μ with $M = 10000$, $\lambda = 1$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

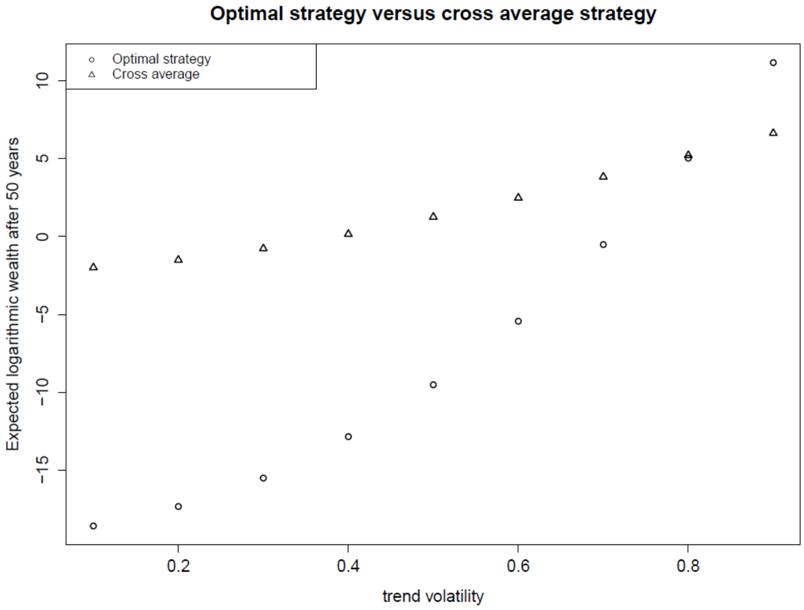

Figure 4.10: The expected logarithmic returns of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days)as functions of σ_μ with $M = 10000$, $\lambda = 2$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

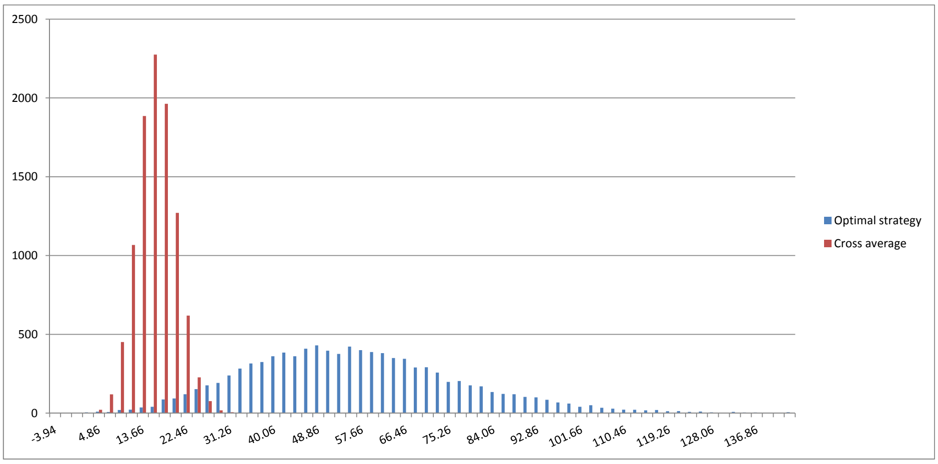

Figure 4.11: Empirical distribution of the logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 90\%$, $\lambda = 1$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

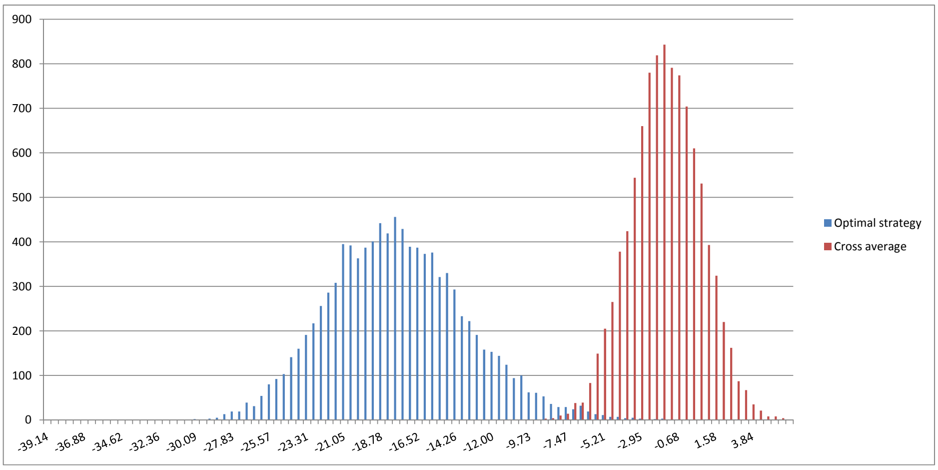

Figure 4.12: Empirical distribution of the expected logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 10\%$, $\lambda = 1$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

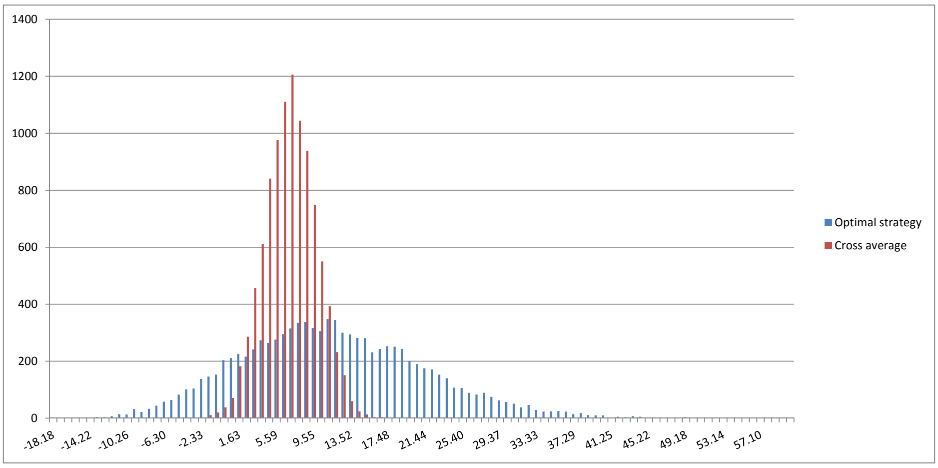

Figure 4.13: Empirical distribution of the expected logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma^a_\mu = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 90\%$, $\lambda = 2$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

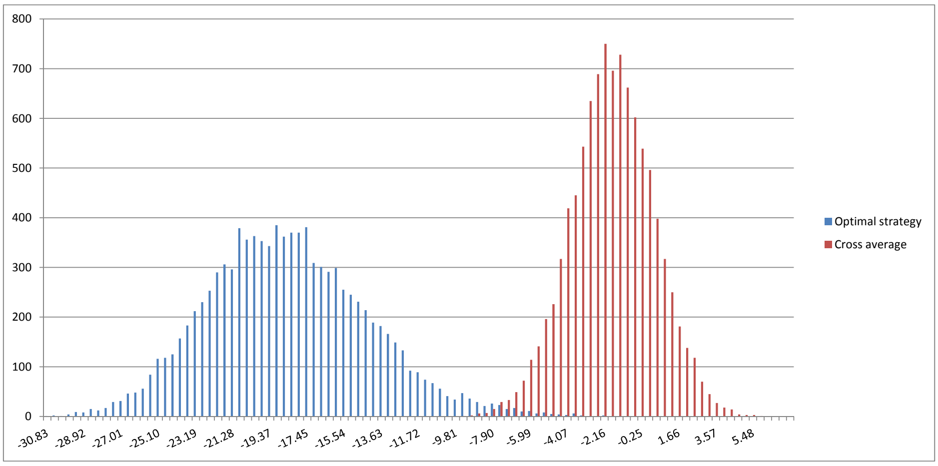

Figure 4.14: Empirical distribution of the expected logarithmic return of the optimal strategy (with $\lambda^a = 1$ and $\sigma_\mu^a = 90\%$) and of the cross average strategy ($L_1 = 5$ days and $L_1 = 252$ days) with $M = 10000$, $\sigma_\mu = 10\%$, $\lambda = 2$, $\alpha = 4$, $\epsilon = 5\%$, $V_\infty = V_0 = 0.3^2$, $\rho = -60\%$ and $T = 50$ years

Tests on real data

Here we test the performances of the two previous strategies on real data. The performance of a strategy is evaluated with the annualised Sharpe ratio indicator (see Sharpe (1966)) on relative daily returns. For the optimal strategy, we assume that $\tau = 252$ business days, that $m = 0.1$ (it has no impact on the Sharpe ratio indicator), and that the volatility σ_S is computed over all the data and used since the beginning of the backtest. For the cross moving average strategy, we keep the same assumptions than the previous section (a window of x days is replaced by a window of x business days). The universe of underlyings are nine stock indexes (the SP 500 Index, the Dow Jones Industrial average Index, the Nasdaq Index, the Euro Stoxx 50 Index, the Cac 40 Index, the Dax Index, the Nikkei 225 Index, the Ftse 100 Index and the Asx 200 Index) and nine forex exchange rates (EUR/CNY, EUR/USD, EUR/JPY, EUR/GBP, EUR/CHF, EUR/MYR, EUR/BRL, EUR/AUD and EUR/ZAR). The period considered is from 12/22/1999 to 2/1/2015. In this test, we assume that these indexes are tradable and that the traded price is given by the closing price of the underlying. The backtest is done without transaction costs. For each strategy, the reallocation is made on a daily frequency. The figure 4.9 gives the measured annualised

Sharpe ratio of the 18 underlyings for each strategy. We observe that the cross moving average strategy outperforms the optimal strategy except for the EUR/BRL.

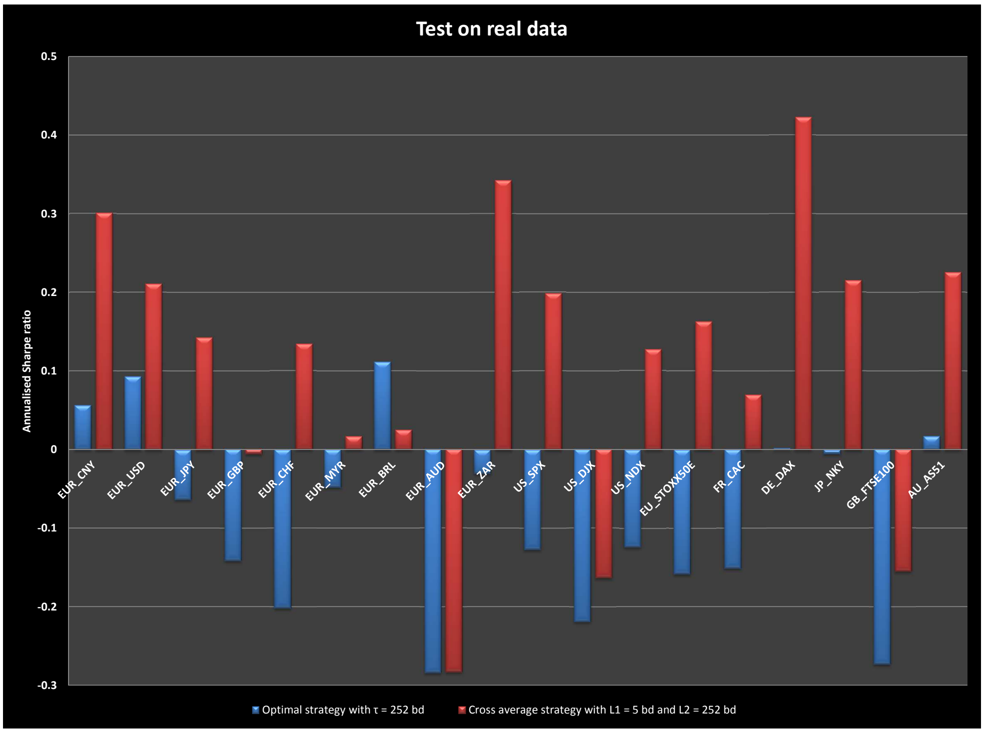

Figure 4.15: Sharpe ratio of the optimal strategy (with $\tau = 252$ bd) and of the cross average strategy ($L_1 = 5$ bd and $L_1 = 252$ bd) on real data from 12/22/1999 to 2/1/2015

4.5 Conclusion

The present work quantifies the performances of the optimal strategy under parameters mis-specification and of a cross moving average strategy using geometric moving averages with a model based on an unobserved mean-reverting diffusion.

For the optimal strategy, we show that the asymptotic expectation of the logarithmic returns is a an increasing function of the signal-to-noise ratio and a decreasing function of the trend mean reversion speed.

We find that, under parameters mis-specification, the performance can be positive under some conditions on the model and strategy parameters. Under the same assumptions, we show the existence of an optimal duration which is equal to the Kalman filter duration if the parameters are well-specified.

For the cross moving average strategy, we also provide the asymptotic logarithmic return of this strategy as a function of the model parameters.

Moreover, the simulations show that, with a model based on an unobserved mean-reverting diffusion, and even with a stochastic volatility, technical analysis investment is more robust than the optimal trading strategy. The empirical tests on real data confirm this conclusion.

Conclusions Générales

L'objectif principal de cette thèse était d'apporter de nouveaux résultats théoriques concernant la notion d'investissement basé sur des modèles stochastiques.
Les résultats du premier papier ont permis de répondre à la faisabilité du problème de la calibration, l'estimation et la détection de la tendance d'un actif. En effet, dans un premier temps, la borne de Cramer Rao et les simulations effectuées ont pu montrer qu'il était impossible d'avoir une bonne précision sur les paramètres en un temps raisonnable. C'est en étudiant ensuite l'impact d'une mauvaise calibration sur l'estimation de la tendance que l'on a montré que l'erreur introduite était non négligeable. Malgré le faible ratio signal-sur-bruit présent dans les données financières, nous avons cependant pu extraire une information statistiquement exploitable : il s'agit de la détection du signe de la tendance. Ce fait est rendu possible en raison de la corrélation non nulle entre la tendance et l'estimateur utilisé, qu'il soit bien ou mal calibré.
Dans le second papier, nous avons quantifié, de façon analytique, l'impact de la non-observabilité de la tendance sur la stratégie de trading optimale. Cette stratégie a été étudiée en supposant que l'on observait ou non la tendance. Dans chacun des cas, nous avons explicité la limite asymptotique de l'espérance et la variance du rendement logarithmique mais également le ratio de Sharpe asymptotique en fonction du ratio signal-sur-bruit et de la vitesse de retour à la moyenne de la tendance. Sous observations partielles du système, nous avons trouvé que le ratio de Sharpe asymptotique n'est pas une fonction monotone de la vitesse de retour à la moyenne de la tendance (croissante puis décroissante) mais qu'il demeure une fonction croissante du ratio signal-sur-bruit comme dans le cas où la tendance est observée. Nous avons conclu cette étude en montrant que le ratio de Sharpe asymptotique de la stratégie optimale avec observations partielles ne peut dépasser $\frac{2}{3^{3/2}} * 100\%$ du ratio de Sharpe asymptotique de la stratégie optimale avec informations complètes.
Le troisième papier étudie la robustesse de la stratégie optimale avec une mauvaise calibration et compare sa performance avec celle d'une stratégie d'analyse technique. Pour ce faire, nous avons caractérisé, de façon analytique, l'espérance asymptotique du rendement logarithmique de chacune de

ces deux stratégies. Nous avons également trouvé quelles conditions devaient être satisfaites sur les paramètres de l'allocation et du modèle pour que la stratégie optimale mal calibrée ait une performance positive. Finalement, nous avons pu exhiber un exemple d'une stratégie d'analyse technique qui est plus robuste que la stratégie optimale.
Comme cette thèse apporte des réponses théoriques à des problèmes pratiques, ce travail est utile à la fois pour le monde académique et pour l'industrie financière.
Finalement, je tiens encore à réitérer mes chaleureux remerciements à toutes les personnes qui m'ont aidé à réaliser cette thèse.

Appendices

Appendix A: discrete Kalman filter

Framework

This section is based on Kalman (1960). The discrete Kalman filter is a recursive method. Consider two objects: the observations $\{y_k\}$ and the states of the system $\{x_k\}$. This filter is based on a Gauss-Markov first order model. Consider the following system:

$$\begin{aligned} x_{k+1} &= F_k x_k + v_k, \\ y_k &= H_k x_k + u_k. \end{aligned}$$

The first equation is an *a priori* model, the transition equation of the system. The matrix F_k is the transition matrix and v_k is the transition noise. The second equation is the measurement equation. The matrix H_k is named the measurement matrix and u_k is the measurement noise. The aim is to identify the underlying process $\{x_k\}$. The two noises are supposed white, Gaussian, centered and decorrelated. In particular:

$$\mathbb{E}\left[\begin{pmatrix} u_k \\ v_k \end{pmatrix}\begin{pmatrix} u_l \\ v_l \end{pmatrix}^T\right] = \begin{pmatrix} R_k^u & 0 \\ 0 & R_k^v \end{pmatrix}\delta_{kl}.$$

The two noises are also supposed independent of x_k and the initial state is Gaussian. So, it can be proved with a recurrence that all states are Gaussian. Therefore, just the mean and the covariance matrix are needed for the characterization of the state. The estimation is given by two steps. The first one is an *a priori* estimation given $\hat{x}_{k+1/k}$ and $\Gamma_{k+1/k} = \mathbb{E}\left[(x_{k+1} - \hat{x}_{k+1/k})(x_{k+1} - \hat{x}_{k+1/k})^T\right]$. When the new observation is available, a correction of the estimation is done to obtain $\hat{x}_{k+1/k+1}$ and $\Gamma_{k+1/k+1} = \mathbb{E}\left[(x_{k+1} - \hat{x}_{k+1/k+1})(x_{k+1} - \hat{x}_{k+1/k+1})^T\right]$. This is the *a posteriori* estimation. The criterion considered for the *a posteriori* estimation is the least squares method, which corresponds to the minimization of the trace of $\Gamma_{k+1/k+1}$.

Filter

The prediction (*a priori* estimation) is given by

$$\begin{aligned}
\hat{x}_{k+1/k} &= F_k \hat{x}_{k/k}, \\
\Gamma_{k+1/k} &= F_k \Gamma_{k/k} F_k^T + R_k^v.
\end{aligned}$$

The *a posteriori* estimation is a correction of the *a priori* estimation. A gain is introduced to do this correction:

$$\hat{x}_{k+1/k+1} = \hat{x}_{k+1/k} + K_{k+1} \left(y_{k+1} - H_{k+1} \hat{x}_{k+1/k} \right).$$

As explained above, the gain K_{k+1} is found by least squares method, which corresponds to

$$\frac{\partial trace \left(\Gamma_{k+1/k+1} \right)}{\partial K_{k+1}} = 0.$$

With the classical lemma of derivation for matrix, the gain is found:

$$\begin{aligned}
K_{k+1} &= \Gamma_{k+1/k} H_{k+1}^T \left[H_{k+1} \Gamma_{k+1/k} H_{k+1}^T + R_{k+1}^u \right]^{-1}, \\
\Gamma_{k+1/k+1} &= \left(I_d - K_{k+1} H_{k+1} \right) \Gamma_{k+1/k}.
\end{aligned}$$

Appendix B: Iterative methods for the inverse and the determinant of the covariance matrix

In this appendix, we provide iterative methods for the inverse and the determinant of the covariance matrix.

Inverse of the covariance matrix

The use of the Matrix Inversion Lemma on Equation (2.10) gives:

$$\Sigma^{-1}_{y_{1:N}|\theta} = \Sigma^{-1}_{\mu_{1:N}|\theta} - \Sigma^{-1}_{\mu_{1:N}|\theta} A_N^{-1} \Sigma^{-1}_{\mu_{1:N}|\theta},$$

where $A_N = \frac{\delta}{\sigma_S^2} I_N + \Sigma^{-1}_{\mu_{1:N}|\theta}$. Then, we have to compute the inverse of the matrices A_N and $\Sigma_{\mu_{1:N}|\theta}$.

Inverse of the matrix A_N

Suppose that A_N^{-1} is computed. The matrix A_{N+1} can be broken into four sub-matrices:

$$A_{N+1} = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix},$$

where

$$\begin{aligned}
B_1 &= \frac{\delta}{\sigma_S^2} + \frac{2\lambda\left(e^{\lambda\delta} + e^{-\lambda\delta}\right)}{\sigma_\mu^2\left(e^{\lambda\delta} - e^{-\lambda\delta}\right)}, \\
B_2 &= \begin{pmatrix} \frac{-2\lambda}{\sigma_\mu^2\left(e^{\lambda\delta} - e^{-\lambda\delta}\right)} & 0 & \cdots & 0 \end{pmatrix}, \\
B_3 &= B_2^T, \\
B_4 &= A_N.
\end{aligned}$$

Therefore, the matrix A_{N+1} can be inverted blockwise.

Inverse of the matrix $\Sigma_{\mu_{1:N}|\theta}$

The following lemma is used (see Akesson & Lehoczky (1998) for details):

Lemma 4.5.1. *Let μ be an Ornstein Uhlenbeck process with parameters*

$\theta = (\lambda, \sigma_\mu)$. *The covariance matrix of* $\mu_1, .., \mu_N$ *is* $\Sigma_{\mu_{1:N}|\theta}$. *Then:*

$$\Sigma^{-1}_{\mu_{1:N}|\theta} = \frac{2\lambda}{\sigma_\mu^2 \left(e^{\lambda\delta} - e^{-\lambda\delta}\right)} B_N,$$

$$B_N = \begin{pmatrix} e^{\lambda\delta}+e^{-\lambda\delta} & -1 & 0 & \cdots & \cdots & 0 \\ -1 & e^{\lambda\delta}+e^{-\lambda\delta} & -1 & \ddots & & \vdots \\ 0 & -1 & e^{\lambda\delta}+e^{-\lambda\delta} & -1 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & -1 & e^{\lambda\delta}+e^{-\lambda\delta} & -1 & 0 \\ \vdots & & \ddots & -1 & e^{\lambda\delta}+e^{-\lambda\delta} & -1 \\ 0 & \cdots & \cdots & 0 & -1 & e^{\lambda\delta} \end{pmatrix}.$$

Therefore, the inverse of the matrix $\Sigma_{\mu_{1:N+1}}$ is given by:

$$\Sigma^{-1}_{\mu_{1:N+1}|\theta} = \begin{pmatrix} \frac{2\lambda\left(e^{\lambda\delta}+e^{-\lambda\delta}\right)}{\sigma_\mu^2\left(e^{\lambda\delta}-e^{-\lambda\delta}\right)} & \frac{-2\lambda}{\sigma_\mu^2\left(e^{\lambda\delta}-e^{-\lambda\delta}\right)} \quad 0 \quad \cdots \quad 0 \\ \frac{-2\lambda}{\sigma_\mu^2\left(e^{\lambda\delta}-e^{-\lambda\delta}\right)} & \\ 0 & \Sigma^{-1}_{\mu_{1:N}|\theta} \\ \vdots & \\ 0 & \end{pmatrix}.$$

Procedure Finally, at time t, the inverse of the covariance matrix is given by the following protocol:

- Computation of the matrix A_t^{-1} using A_{t-1}^{-1}.
- Computation of the matrix $\Sigma^{-1}_{\mu_{1:t}|\theta}$ using $\Sigma^{-1}_{\mu_{1:t-1}|\theta}$.
- Using $\Sigma^{-1}_{y_{1:t}|\theta} = \Sigma^{-1}_{\mu_{1:t}|\theta} - \Sigma^{-1}_{\mu_{1:t}|\theta} A_t^{-1} \Sigma^{-1}_{\mu_{1:t}|\theta}$, the matrix $\Sigma^{-1}_{y_{1:t}|\theta}$ is obtained.

Determinant of the covariance matrix

The iterative computation of $\det\left(\Sigma_{y_{1:N}|\theta}\right)$ is based on the following lemma:

Lemma 4.5.2. *The determinant of the matrix* $\Sigma_{y_{1:N}|\theta}$ *is given by:*

$$\det\left(\Sigma_{y_{1:N}|\theta}\right) = \frac{\det\left(I_N + \frac{\sigma_S^2}{\delta}\Sigma^{-1}_{\mu_{1:N}|\theta}\right)}{\det\left(\Sigma^{-1}_{\mu_{1:N}|\theta}\right)}, \tag{4.39}$$

and for $N \geq 2$, we have:

$$\begin{aligned}\det\left(\Sigma^{-1}_{\mu_{1:N+1}|\theta}\right) &= g(\lambda,\sigma_\mu)\left(e^{\lambda\delta}+e^{-\lambda\delta}\right)\det\left(\Sigma^{-1}_{\mu_{1:N}|\theta}\right) \\ &\quad -g(\lambda,\sigma_\mu)^2\det\left(\Sigma^{-1}_{\mu_{1:N-1}|\theta}\right), \\ \det\left(I_{N+1}+\frac{\sigma_S^2}{\delta}\Sigma^{-1}_{\mu_{1:N+1}|\theta}\right) &= \left(1+\frac{\sigma_S^2}{\delta}g(\lambda,\sigma_\mu)\left(e^{\lambda\delta}+e^{-\lambda\delta}\right)\right)\det\left(I_N+\frac{\sigma_S^2}{\delta}\Sigma^{-1}_{\mu_{1:N}|\theta}\right) \\ &\quad -\left(\frac{\sigma_S^2}{\delta}g(\lambda,\sigma_\mu)\right)^2\det\left(I_{N-1}+\frac{\sigma_S^2}{\delta}\Sigma^{-1}_{\mu_{1:N-1}|\theta}\right),\end{aligned}$$

where

$$g\left(\lambda,\sigma_\mu\right)=\frac{2\lambda}{\sigma_\mu^2\left(e^{\lambda\delta}-e^{-\lambda\delta}\right)}.$$

Proof. The multiplication of Equation (2.10) by $\Sigma^{-1}_{\mu_{1:N}|\theta}$ gives:

$$\Sigma^{-1}_{\mu_{1:N}|\theta}\Sigma_{y_{1:N}|\theta} = I_N+\frac{\sigma_S^2}{\delta}\Sigma^{-1}_{\mu_{1:N}|\theta}.$$

Equation (4.39) follows. Using Lemma 4.5.1, The matrices $\left(I_N+\frac{\sigma_S^2}{\delta}\Sigma^{-1}_{\mu_{1:N}|\theta}\right)$ and $\Sigma^{-1}_{\mu_{1:N}|\theta}$ are tridiagonal. The recursive computation of their determinant is then possible. □

Appendix C: Proof of Proposition 18

Proof. Let K be a $(\mathbb{P}, \{\mathcal{F}_t\})$ martingale defined by:

$$\frac{dK_t}{K_t} = \frac{-\mu_t}{\sigma_S} dW_t^S,$$

and the probability measure $\widetilde{\mathbb{P}}$ defined by:

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = K_T.$$

With the Girsanov's theorem, it follows that the process:

$$\widetilde{W}_t^S = W_t^S + \int_0^t \frac{\mu_s}{\sigma_S} ds,$$

is a $\left(\widetilde{\mathbb{P}}, \{\mathcal{F}_t\}\right)$ Wiener process. Note also that:

$$\frac{dS_t}{S_t} = \sigma d\widetilde{W}_t^S.$$

Now, introduce the process N, defined by:

$$N_t = \widetilde{W}_t^S - \int_0^t \frac{E\left[\mu_s | \mathcal{F}_s^S\right]}{\sigma_S} ds,$$

as $\widetilde{W}_t^S$ and $\mathbb{E}\left[\mu_t | \mathbb{F}_t^S\right]$ are $\left\{\mathcal{F}_t^S\right\}$ measurable, N_t is $\left\{\mathcal{F}_t^S\right\}$ measurable. The process N is also integrable. Let τ be a bounded stopping time. we have

$$\mathbb{E}\left[N_\tau\right] = \mathbb{E}\left[N_0\right] = 0.$$

Then, N is a continuous martingale and $N_0 = 0$. Note that

$$d\left\langle N \right\rangle_t = dt$$

Using Levy's criteria, the process N is a $\left(\mathbb{P}, \left\{\mathcal{F}_t^S\right\}\right)$ Wiener process. □

Appendix D: Auto-covariance function of the square steady state Kalman filter

the following lemma gives the auto-covariance function of the process $\hat{\mu}^2$:

Lemma 4.5.3. *Consider the process $\hat{\mu}$ defined in Equation (3.4). Its auto-covariance function is given by:*

$$\mathbb{C}ov\left[\hat{\mu}_s^2, \hat{\mu}_t^2\right] = \frac{\lambda^2\sigma_S^4\left(\beta-1\right)^4}{2}e^{-2\lambda t}\left(e^{2\lambda s}+e^{-2\lambda s}-2\right), \tag{4.40}$$

with $0 \le s \le t$.

Proof. Since $\hat{\mu}$ is a centred Ornstein Uhlenbeck process, there exists a Brownian motion B such that, for all $s \in \mathbb{R}_+$:

$$\hat{\mu}_s = e^{-\lambda s}\lambda\sigma_S\left(\beta-1\right)B_{f(s)},$$

where $f\left(s\right) = \frac{e^{2\lambda s}-1}{2\lambda}$ is a time change. Then, for all s, t such that $0 \le s \le t$, we have:

$$\mathbb{C}\mathrm{ov}\left[\hat{\mu}_s^2, \hat{\mu}_t^2\right] = e^{-2\lambda(t+s)}\lambda^4\sigma_S^4\left(\beta-1\right)^4\mathbb{C}\mathrm{ov}\left[B_{f(s)}^2, B_{f(t)}^2\right].$$

Since B is a Wiener process:

$$\mathbb{E}\left[B_{f(s)}^2\right] = f\left(s\right).$$

Let $\left\{\mathcal{F}_t^B\right\}$ be the filtration generated by the process B. So:

$$\begin{aligned}
\mathbb{E}\left[B_{f(s)}^2, B_{f(t)}^2\right] &= \mathbb{E}\left[B_{f(s)}^2\mathbb{E}\left[B_{f(t)}^2|\mathcal{F}_s^B\right]\right] \\
&= \mathbb{E}\left[B_{f(s)}^2\mathbb{E}\left[\left(B_{f(s)}^2+2\int_{f(s)}^{f(t)}B_u dB_u+f(t)-f(s)\right)|\mathcal{F}_s^B\right]\right] \\
&= \mathbb{E}\left[B_{f(s)}^2\left(B_{f(s)}^2+f\left(t\right)-f\left(s\right)\right)\right] \\
&= 3f\left(s\right)^2+\left(f\left(t\right)-f\left(s\right)\right)f\left(s\right).
\end{aligned}$$

Then

$$\mathbb{E}\left[B_{f(s)}^2, B_{f(t)}^2\right] = 2f\left(s\right)^2+f\left(t\right)f\left(s\right),$$

and Equation (4.40) follows using:

$$\mathbb{C}\mathrm{ov}\left[B_{f(s)}^2, B_{f(t)}^2\right] = \mathbb{E}\left[B_{f(s)}^2, B_{f(t)}^2\right]-\mathbb{E}\left[B_{f(s)}^2\right]\mathbb{E}\left[B_{f(t)}^2\right].$$

□

Appendix E: Proof of Proposition 23

Proof. Let K be a $(\mathbb{P}, \{\mathcal{F}_t\})$ martingale defined by:

$$\frac{dK_t}{K_t} = \frac{-\mu_t}{\sigma_S} dW_t^S,$$

and the probability measure $\widetilde{\mathbb{P}}$ defined by:

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = K_T.$$

With the Girsanov's theorem, it follows that the process:

$$\widetilde{W}_t^S = W_t^S + \int_0^t \frac{\mu_s}{\sigma_S} ds,$$

is a $\left(\widetilde{\mathbb{P}}, \{\mathcal{F}_t\}\right)$ Wiener process. Note also that:

$$\frac{dS_t}{S_t} = \sigma d\widetilde{W}_t^S.$$

Now, introduce the process N, defined by:

$$N_t = \widetilde{W}_t^S - \int_0^t \frac{E\left[\mu_s | \mathcal{F}_s^S\right]}{\sigma_S} ds,$$

as $\widetilde{W}_t^S$ and $\mathbb{E}\left[\mu_t | \mathbb{F}_t^S\right]$ are $\left\{\mathcal{F}_t^S\right\}$ measurable, N_t is $\left\{\mathcal{F}_t^S\right\}$ measurable. The process N is also integrable. Let τ be a bounded stopping time. we have

$$\mathbb{E}\left[N_\tau\right] = \mathbb{E}\left[N_0\right] = 0$$

N is a continuous martingale and $N_0 = 0$. Note that

$$d\left\langle N\right\rangle_t = dt$$

Using Levy criteria, the process N is a $\left(\mathbb{P}, \left\{\mathcal{F}_t^S\right\}\right)$ Wiener process. □

Bibliography

Akesson, F., & Lehoczky, J. 1998. *Discrete Eigenfunction Expansion of Multi-Dimensional Brownian Motion and the Ornstein-Uhlenbeck Process*. Tech. rept. preprint.

Anane, M., & Abergel, F. 2015. Empirical Evidence of Market Inefficiency: Predicting Single-Stock Returns. *Pages 3–66 of:* Abergel, F., Aoyama, H., Chakrabarti, B. K., Chakraborti, A., & Ghosh, A. (eds), *Econophysics and Data Driven Modelling of Market Dynamics*. New Economic Windows. Springer International Publishing.

Bachelier, L. 1900. *Théorie de la Spéculation*. Ph.D. thesis, Sorbonne.

Benmiloud, B., & Pieczynski, W. 1995. *Parameter Estimation in Hidden Markov Chains and Segmentation of Images*. Tech. rept. Institut National des Télécommunications.

Bjork, T., Davis, Mark H.A., & Landén, C. 2010. *Optimal Investment under Partial Information*. SSE/EFI Working Paper Series in Economics and Finance 739. Stockholm School of Economics.

Black, F., & Scholes, M. 1973. The pricing of options and corporate liabilities. *Journal of political economy*.

Blanchet-Scalliet, Christophette, Diop, Awa, Gibson Brandon, Rajna, Talay, Denis, & Tanré, Etienne. 2007. Technical Analysis Compared to Mathematical Models Based Methods Under Parameters Mis-specification. *Journal of Banking and Finance*, **31**(5), 1351–1373.

Brendle, S. 2006a. Portfolio selection under incomplete information. *Stochastic Processes and their Applications*, **116**(5), 701 – 723.

Brendle, S. 2006b. Portfolio selection under incomplete information. *Stochastic Processes and their applications*.

Brockwell, Peter J., & Davis, Richard A. 2002a. *Introduction to time series and forecasting*. Springer-Verlag New York.

Brockwell, Peter J., & Davis, Richard A. 2002b. *Time series: Theory and Methods*. Springer-Verlag New York.

Brown, David P, & Jennings, Robert H. 1989. On technical analysis. *Review of Financial Studies*, **2**(4), 527–551.

Bruder, B., & Gaussel, N. 2011. *Risk-Return Analysis of Dynamic Investment Strategies*. Tech. rept. Lyxor.

Casarin, R., & Marin, J.M. 2007. *Online data processing: comparison of Bayesian regularized particle filters*. Tech. rept. INRIA.

Challet, D., & Bel Hadj Ayed, A. 2015. Do Google Trend data contain more predictability than price returns? *Journal of investment strategies*.

Chib, S., Nardari, F., & Shephard, N. 2002. Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, June.

Clauss, P. 2011. *Gestion de portefeuille - Une approche quantitative*. Dunod.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. 1985. A theory of the term structure of interest rates. *Econometrica*.

D. Duffie, W. Zame. 1989. The consumption-based capital asset pricing model. *Econometrica*.

Dahia, K. 2005 (Jan.). *Nouvelles méthodes en filtrage particulaire. Application au recalage de navigation inertielle par mesures altimétriques*. Ph.D. thesis, Université Joseph Fourier.

D.L. Ocone, I. Karatzas. 1991. A generalized Clark representation formula, with application to optimal portfolios. *Stochastics and Stochastic Reports*.

Edwards, Robert D, Magee, John, & Bassetti, WHC. 2007. *Technical analysis of stock trends*. CRC Press.

Eraker, B. 1998. MCMC Analysis of Diffusion Models with Application to Finance. *Journal of Business and Economic Statistics*.

Fama, E. F. 1969. Efficient Capital Markets : A Review of Theory and Empirical Work. *Journal of finance*, **25**(2), 383–417.

Genon-Catalot, V. 2009. *Parameter estimation for stochastic differential equations from noisy observations. Maximum likelihood and filtering techniques*. Lipari Biomathematics summer school.

H. He, N. Pearson. 1991. Consumption and portfolio policies with incomplete markets: the infinite dimensional case. *J. Economic Theory*.

Heston, S. L. 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*.

I. Karatzas, J.P. Lehoczky, S.E. Shreve. 1987. Optimal portfolio and consumption decisions for a “small investor” on a finite horizon. *SIAM J. Control Optim.*

Isserlis, L. 1918. On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*.

Jacquier, E., Polson, N. G., & Rossi, P. E. 1994. *Bayesian Analysis of Stochastic Volatility Models*.

J.C. Cox, C.F. Huang. 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *J. Economic Theory*.

JJ.C. Cox, J.E. Ingersoll, S.A. Ross. 1985. An intertemporal general equilibrium model of asset prices. *Econometrica*.

Kalman, R.E. 1960. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*.

Kalman, R.E., Engiar, T.S., & bucy, R.S. 1962. *Fundamental study of adaptive control systems*. Tech. rept. DTIC Document.

Karatzas, I., & Zhao, X. 2001. Bayesian adaptive portfolio optimization. *Handbook of Mathematical Finance, Optimization, Pricing, Interest Rates, and Risk Management*.

Kim, S., Shephard, N., & Chib, S. 1998. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *The Review of Economic Studies*, 361–393.

Lakner, P. 1995. Utility maximization with partial information. *Stochastic Processes Appl.*

Lakner, P. 1998. Optimal trading strategy for an investor: the case of partial information. *Stochastic Processes and their Applications*.

Laskry, J.M., & Lions, P.L. 1999. Contrôle stochastique avec informations partielles et applications à la Finance. *Comptes Rendus de l’Académie des Sciences - Series I - Mathematics*, **328**(11), 1003 – 1010.

Lempérière, Y., Deremble, C., Seager, P., Potters, M., & Bouchaud, J. P. 2014. Two centuries of trend following. *Journal of Investment Strategies*.

Leroux, Brian G. 1992. Maximum-likelihood estimation for hidden Markov models. *Stochastic Processes and their Applications*, **40**(1), 127 – 143.

Liptser, R. S., & Shiriaev, A. N. 1977. *Statistics of random processes I*. Springer-Verlag New York.

Lord, R., Koekkoek, R., & Dijk, D. V. 2010. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, **10**(2), 177–194.

Markovitz, H. 1952. Portfolio Selection. *The Journal of Finance*.

Merton, R. C. 1990. *Continuouse-Time Finance*. Blackwell.

Mikhailov, S., & Nögel, U. 2003. Heston's stochastic volatility model-implementation, calibration and some extensions. *Wilmott magazine*.

Monahan, John F. 1983. Fully Bayesian analysis of ARMA time series models. *Journal of Econometrics*, **21**(3), 307 – 331.

O. Cappé, T. Rydén, & Moulines, E. 2005. *Inference in Hidden Markov Models*. Springer-Verlag.

Pham, H., & Quenez, M.C. 2001. Optimal Portfolio in Partially Observed Stochastic Volatility Models. *Annals of Applied Probability*, **11**, 210–238.

Rieder, U., & Bauerle, N. 2005. Portfolio optimization with unobservable Markov-modulated drift process. *Journal of Applied Probability*.

Sass, J., & Haussmann, U.G. 2004. Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain. *Finance and Stochastics*.

Schweppe, F.C. 1965. Evaluation of likelihood functions for Gaussian signals. *IEEE Transactions on Information Theory*.

Schwert, G. W. 2003. Anomalies and market efficiency. *In:* Constantinides, G.M., Harris, M., & Stulz, R. M. (eds), *Handbook of the Economics of Finance*, vol. 1. Elsevier.

Sharpe, W. F. 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*.

Sharpe, W. F. 1966. Mutual Fund Performance. *The Journal of Business*.

Tamaki, K. 2008. The Bernstein-von Mises Theorem for Stationary Processes. *Journal of the Japan statistical society*, **38**(2), 311–323.

Taylor, Mark P, & Allen, Helen. 1992. The use of technical analysis in the foreign exchange market. *Journal of international Money and Finance*, **11**(3), 304–314.

Tideman, TN. 1972. Basic Economic Principles.

Treynor, J. L. 1961. Market Value, Time, and Risk. *Independent.*

Whittle, P. 1953. The analysis of multiple stationary time series. *Journal of the Royal Statistical Society. Series B (Methodological)*, **15**(1), 125–139.

Zhu, Yingzi, & Zhou, Guofu. 2009. Technical analysis: An asset allocation perspective on the use of moving averages. *Journal of Financial Economics.*