Allocation stratégique d'actifs et ALM pour les régimes de retraite

Alaeddine Faleh

▶ **To cite this version:**

Alaeddine Faleh. Allocation stratégique d'actifs et ALM pour les régimes de retraite. Gestion de portefeuilles [q-fin.PM]. Université Claude Bernard - Lyon I, 2011. Français. NNT : . tel-00689907v1

HAL Id: tel-00689907

https://theses.hal.science/tel-00689907v1

Submitted on 20 Apr 2012 (v1), last revised 11 Sep 2012 (v4)

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

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

N° d'ordre : 84-2011 Année 2011

THÈSE
présentée
devant l'UNIVERSITÉ CLAUDE BERNARD - LYON 1 – ISFA
pour l'obtention du
DIPLÔME DE DOCTORAT
Spécialité sciences actuarielle et financière

présentée et soutenue publiquement le 13/05/2011
par Mr. Alaeddine FALEH

Allocation stratégique d'actifs et ALM pour les régimes de retraite

Directeur de thèse : Pr. Frédéric PLANCHET
Dr. Didier RULLIERE (co-directeur)

Composition du jury :

Président : Pr. Jean-Claude AUGROS
Professeur à l'Université Claude Bernard Lyon1 (I.S.F.A.)

Rapporteurs : Pr. Jean-Paul LAURENT
Professeur à l'Université Paris 1 Panthéon-Sorbonne

Pr. François DUFRESNE
Professeur à l'Université de Lausanne (Suisse)

Directeur de thèse : Pr. Frédéric PLANCHET
Professeur associé à l'Université Claude Bernard Lyon1 (I.S.F.A.)

Dr. Didier RULLIERE
Maître de conférences à l'Université Claude Bernard Lyon1 (I.S.F.A.)

Suffragant : Pr. Jean-Pierre AUBIN
Professeur à l'Ecole Polytechnique de Paris

Dr. Guillaume LEZAN
Caisse des Dépôts et Consignations (CDC)

-Ecole Doctorale Sciences Economiques et de Gestion-
-Institut de Science Financière et d'Assurances (Univ. Lyon I) -
-Laboratoire de Sciences Actuarielle et Financière (équipe d'accueil n°2429)-

RÉSUMÉ

La présente thèse s'intéresse aux modèles d'allocation stratégiques d'actifs et à leurs applications pour la gestion des réserves financières des régimes de retraite par répartition, en particulier ceux partiellement provisionnés. L'étude de l'utilité des réserves pour un système par répartition et *a fortiori* de leur gestion reste un sujet peu exploré. Les hypothèses classiques sont parfois jugées trop restrictives pour décrire l'évolution complexe des réserves. De nouveaux modèles et de nouveaux résultats sont développés à trois niveaux : la génération de scénarios économiques (GSE), les techniques d'optimisation numérique et le choix de l'allocation stratégique optimale dans un contexte de gestion actif-passif (ALM).

Dans le cadre de la génération de scénarios économiques et financiers, certains indicateurs de mesure de performance du GSE ont été étudiés. Par ailleurs, des améliorations par rapport à ce qui se pratique usuellement lors de la construction du GSE ont été apportées, notamment au niveau du choix de la matrice de corrélation entre les variables modélisées. Concernant le calibrage du GSE, un ensemble d'outils permettant l'estimation de ses différents paramètres a été présenté.

Cette thèse a également accordé une attention particulière aux techniques numériques de recherche de l'optimum, qui demeurent des questions essentielles pour la mise en place d'un modèle d'allocation. Une réflexion sur un algorithme d'optimisation globale d'une fonction non convexe et bruitée a été développée. L'algorithme permet de moduler facilement, au moyen de deux paramètres, la réitération de tirages dans un voisinage des points solutions découverts, ou à l'inverse l'exploration de la fonction dans des zones encore peu explorées.

Nous présentons ensuite des techniques novatrices d'ALM basées sur la programmation stochastique. Leur application a été développée pour le choix de l'allocation stratégique d'actifs des régimes de retraite par répartition partiellement provisionnés. Une nouvelle méthodologie pour la génération de l'arbre des scénarios a été adoptée à ce niveau. Enfin, une étude comparative du modèle d'ALM développé avec celui basé sur la stratégie *Fixed-Mix* a été effectuée. Différents tests de sensibilité ont été par ailleurs mis en place pour mesurer l'impact du changement de certaines variables clés d'entrée sur les résultats produits par notre modèle d'ALM.

MOTS-CLÉS

Gestion actif-passif, Allocation d'actifs, Génération de scénarios, Optimisation stochastique, Programmation stochastique, Régime de retraite

ABSTRACT

This thesis focuses on the strategic asset allocation models and on their application for the financial reserve management of a pay-as-you-go (PAYG) retirement schemes, especially those with partial provision. The study of the reserve utility for a PAYG system and of their management still leaves a lot to be explored. Classical hypothesis are usually considered too restrictive for the description of the complex reserve evolution. New models and new results have been developed over three levels : economic scenario generation (ESG), numerical optimization techniques and the choice of optimal strategic asset allocation in the case of an Asset-Liability Management (ALM).

For the generation of financial and economic scenarios, some ESG performance indicators have been studied. Also, we detailed and proposed to improve ESG construction, notably the choice of the correlation matrix between modelled variables. Then, a set of tools were presented so that we could estimate ESG parameters variety.

This thesis has also paid particular attention to numerical techniques of optimum research, which is an important step for the asset allocation implementation. We developed a reflexion about a global optimisation algorithm of a non convex and a noisy function. The algorithm allows for simple modulating, through two parameters, the reiteration of evaluations at an observed point or the exploration of the noisy function at a new unobserved point.

Then, we presented new ALM techniques based on stochastic programming. An application to the strategic asset allocation of a retirement scheme with partial provision is developed. A specific methodology for the scenario tree generation was proposed at this level. Finally, a comparative study between proposed ALM model and Fixed-Mix strategy based model was achieved. We also made a variety of a sensitivity tests to detect the impact of the input values changes on the output results, provided by our ALM model.

KEY-WORDS

Asset-Liability Management, Asset Allocation, Scenario Generation, Stochastic Optimization, Stochastic Programming, Retirement Scheme

REMERCIEMENTS

Je souhaite tout d'abord remercier vivement Prof. Frédéric PLANCHET pour avoir dirigé ma thèse. Je suis particulièrement reconnaissant de la confiance qu'il a su m'accorder dès les premiers kilomètres de ce marathon, ainsi que de ses précieux conseils.

Je veux également remercier Dr. Didier RULLIERE, qui a co-dirigé cette thèse, et dont les suggestions ont été particulièrement profitables au développement de cette thèse.

J'adresse chaleureusement mes remerciements aux membres du jury auxquels je suis reconnaissant pour l'honneur qu'ils me font de juger ce travail.

Cette thèse a également reçu un soutien considérable, à son commencement, de la part du Prof. Jean Claude AUGROS. Je le remercie d'avoir été toujours disponible pour répondre à mes questions.

Mes vifs remerciements s'adressent à Mme. Edith JOUSSEAUME (directrice du département des Investissements et de Comptabilité à la Caisse des Dépôts et Consignation) pour m'avoir accordée sa confiance et son encouragement : des atouts qui ont permis le déroulement de mon travail, au sein de son département, dans les meilleures conditions.

Ma profonde gratitude à Dr. Guillaume LEZAN, mon responsable scientifique à la Caisse des Dépôts et Consignation, qui a passé bien des heures dans la relecture de cette thèse tout en l'enrichissant avec ses pertinentes suggestions.

Cette thèse a bénéficié d'un précieux soutien, à son commencement, de la part de Mr. Pierre LEYGUE (Caisse des Dépôts et consignations) envers qui je reste toujours reconnaissant.

J'adresse un remerciement tout particulier à ma chère femme Asma pour la patience dont elle a fait preuve pendant ces dernières années de travail et l'exploit de me supporter durant cette période.

Merci également à ma famille pour avoir consenti l'éloignement et bien d'autres sacrifices en vue de l'aboutissement de cette thèse.

Finalement, je remercie tous ceux qui ont participé, de près ou de loin, à l'accomplissement de ce travail.

A mes deux chers parents

C'est que toute œuvre scientifique "achevée" n'a d'autre sens que celui de faire naître de nouvelles "questions"…

[Max Weber - *"Le savant et le politique"*]

Sommaire

Sommaire…….....1

Introduction générale……4

Partie I : Les générateurs de scénarios économiques (GSE)……...…..17

Introduction ……..…….…….....18

Chapitre 1 : Etude des GSE…….……..19

I- Présentation théorique ….…….. ……..……..……19

I-1 Problèmes théoriques de modélisation liés aux GSE….……..……...19
I-2 Littérature sur les structures des GSE…….…..…….……24

II- Mise en œuvre d'un GSE…….……..……...30

II-1 Structure schématique de projection de scénarios pour un GSE……..……..…….30
II-2 Méthodologies pour la génération de scénarios économiques……..…….…….34
II-3 Probabilité risque-neutre Vs probabilité réelle…….……..36
II-4 L'exclusion de valeurs négatives pour les modèles de taux…….……..37

III- Mesure de qualité…….…….……..38

III-1 Mesure qualitative de la qualité d'un GSE……38
III-2 Mesure quantitative de la qualité d'un GSE……..40
III-3 Application numérique...…….…….…….....45

Conclusion du chapitre 1……...49

Chapitre 2 : Construction d'un générateur de scénarios économiques…….……50

I- Conception et composantes …….…….…..……50

I-1 Conception théorique …….……..…..…….……...50
I-2 Composantes ……....……..…….……...53

II- Calibrage…… ……..……73

II-1 Présentation du contexte de l'étude …...……..……..……...73
II-2 Calibrage du modèle de l'inflation……..……..……..……...76

II-3 Calibrage du modèle des taux réels...……..……..……79
II-4 Calibrage du modèle des actions..……..……83
II-5 Calibrage du modèle de l'immobilier……..…….....89
II-6 Calibrage du modèle de crédit....……...91

III- Projection des scénarios ……..……93

III-1 Éléments sur la mise en œuvre…….…….…….……..93
III-2 Tests du modèle : paramètres et matrice de corrélation……...……....96
III-3 Résultats de projection……..……..…….....99

Conclusion du chapitre 2…….110

Conclusion de la partie I……...110

Partie II : L'allocation stratégique d'actifs dans le cadre de la gestion actif-passif (ALM)……...……..112

Introduction……...113

Chapitre 1 : Les modèles d'ALM classiques (déterministes)……118

I- Immunisation du portefeuille…..…….118

I-1 Adossement des flux de trésorerie……..……...119
I-2 Adossement par la duration……..…….120

II- Modèles basés sur le surplus…….…….125

II-1 Modèle de Kim et Santomero [1988]...……..……..125
II-2 Modèle de Sharpe et Tint [1990] ……..……...128
II-3 Modèle de Leibowitz [1992]…….……...130
II-4 Duration du Surplus…….……137

Conclusion du chapitre 1…….139

Chapitre 2 : Allocation stratégique d'actifs et stratégie *Fixed-Mix*…….....……140

I- Etude de l'allocation d'actifs dans le cadre de la stratégie *Fixed-Mix*…...……..143

I-1 Présentation…..……. ..143
I-2 Application…...……143

II- Proposition de méthodes numériques ……...……....168

II-1 Introduction…….…….168
II-2 Une grille à pas variable par subdivision systématique……...173
II-3 Une grille à pas variable avec retirage possible…….……..188
II-4 Applications numériques….……191
II-5 Conclusion……...….…....211

Conclusion du chapitre 2…….……212

Chapitre 3 : Allocation stratégique d'actifs et modèles d'ALM dynamique……..…….213

I - L'allocation stratégique d'actifs dans le cadre des modèles classiques d'ALM dynamique…….…….214

I-1 Présentation générale……...……. 214
I-2 Techniques d'assurance de portefeuille …….……...215
I-3 Programmation dynamique……...……...……...…..223

II- L'allocation stratégique d'actifs dans le cadre de la programmation stochastique …….……....227

II-1 Introduction ……...……..……..……..227
II-2 Présentation de la programmation stochastique...……...…….228
II-3 Illustration de la programmation stochastique avec recours dans le cas de la planification de la production…….……..230
II-4 Formulation mathématique de la programmation stochastique avec recours…..236
II-5 Résolution des programmes stochastiques…….……...…...……240

III- Nouvelle approche d'ALM par la discrétisation des scénarios économiques……..…….……....242

III-1 Méthode des quantiles de référence pour le GSE..……...……..242
III-2 Modèle d'optimisation basé sur la programmation stochastique.....…...……247
III-3 Discussion des résultats……. …….……...……253

Conclusion du chapitre 3…….260

Conclusion de la partie II ……260

Conclusion générale…….261

Index des graphiques……...…267

Bibliographie…….....271

Table des matières……...283

INTRODUCTION GÉNÉRALE

Objectifs et enjeux de la thèse

Le champ d'étude offert par les régimes de retraite est très étendu et peut être segmenté selon différentes clés telles que notamment : régimes publics *versus* régimes d'entreprises, gestion en répartition[1] *versus* gestion en capitalisation[2], régimes à cotisations définies *versus* régimes à prestations définies, etc… Trois situations peuvent être recensées pour un régime donné : régime provisionné (cas où le régime est tenu à la couverture totale des engagements souscrits par ses cotisants et retraités actuels), régime partiellement provisionné (couverture d'une partie seulement des engagements souscrits par ses cotisants et retraités actuels) et enfin régime non provisionné.

Pour apprécier la solidité prudentielle d'un régime de retraite, les experts se sont longtemps limités à une approche binaire, selon laquelle il convenait de tout provisionner dans le cas d'un système par capitalisation et rien dans le cas d'un système par répartition. Des travaux récents mettent en évidence l'utilité de réserves pour un système par répartition (cf. Delarue [2001]).

Le fonctionnement financier des régimes de retraite par répartition, en particulier en France, peut être schématisé comme suit : les flux de cotisations permettent de régler les flux de prestations, ensuite le surplus permet, le cas échéant, d'alimenter une réserve destinée à régler une partie des prestations futures. Cette même réserve peut se voir prélever le solde technique débiteur s'il s'avère que les cotisations sont insuffisantes pour régler les prestations. Dans le même temps, la réserve est placée sur les marchés financiers et allouée selon différentes classes d'actifs.

La gestion actif-passif d'un régime par répartition partiellement provisionné peut être basée sur l'optimisation de la valeur de la réserve, compte tenu des contraintes liées au passif qu'il doit respecter. C'est dans ce cadre que le choix d'une « bonne » allocation stratégique joue un rôle essentiel dans le pilotage actif-passif d'un régime de retraite, étant donné que les réserves contribuent à part entière à la solidité du régime.

Cependant une difficulté consiste à définir la « meilleure » stratégie de placement de ces réserves sur les marchés financiers, notamment dans un contexte de fortes incertitudes économiques, comme c'est actuellement le cas du fait de la récente crise financière.

La présente thèse est consacrée aux modèles d'allocation stratégiques d'actifs et à leurs applications pour la gestion des réserves financières des régimes de retraites sur le long-terme.

[1] Mode d'organisation des systèmes de retraite fondé sur la solidarité entre générations. Les cotisations versées par les actifs au titre de l'assurance vieillesse servent immédiatement à payer les retraites. L'équilibre financier des systèmes de retraite par répartition est fonction du rapport entre le nombre de cotisant (population active, taux de croissance des revenus) et celui des retraités. Le système français de retraite est fondé sur le principe de la répartition.

[2] Mode d'organisation des systèmes de retraite dans lequel les cotisations d'un assuré sont placées à son nom durant sa vie active (placements financiers et immobiliers, dont le rendement varie en fonction des taux d'intérêt) avant de lui être restituées sous forme de rente après l'arrêt de son activité professionnelle. La constitution du capital peut s'effectuer à titre individuel ou dans un cadre collectif (accord d'entreprise). En France, seuls les systèmes de retraite dits sur-complémentaires (ex. PREFON, COREM, le PERP ou plan d'épargne populaire) fonctionnent selon le principe de la capitalisation à l'exception du RAFP.

Deux points particuliers font l'objet d'un examen approfondi : d'une part les générateurs de scénarios économiques (GSE) utilisés pour la modélisation des fluctuations des actifs financiers sur le long-terme puis, dans un deuxième temps, les modèles proposés pour l'élaboration de l'allocation d'actifs elle-même. A chaque fois les aspects théoriques précèderont la mise en œuvre opérationnelle sur des exemples.

Concernant les GSE, nous mettons en évidence leurs principales composantes, que ce soit au niveau de la conception théorique ou à celui de la mise en œuvre. Le choix de ces composantes est lié à la vocation finale du générateur de scénarios économiques, que ce soit en tant qu'outil d'évaluation des produits financiers (*pricing*) ou en tant qu'outil de projection et de gestion des risques. Par ailleurs, nous développons une étude sur certains indicateurs de mesure de la performance du GSE comme un outil en amont du processus de prise de décision : la stabilité et l'absence de biais. Une application numérique permettant d'illustrer ces différents points est présentée.

Notre objectif final est de proposer un GSE adapté aux spécificités de la gestion actif-passif d'un régime de retraite. Différents modèles de diffusion et d'estimation de leurs paramètres ont ainsi été envisagés pour le panel des variables financières et macro-économiques qui interviennent dans ce contexte. Cette partie est illustrée par une comparaison des résultats du générateur de scénario proposé avec ceux obtenus avec le modèle d'Ahlgrim et al. [2005].

A ce niveau, notre étude se distingue par l'adoption d'une démarche novatrice dont l'objectif est de tenir compte de l'instabilité temporelle de la corrélation entre les rentabilités des actifs. Autrement dit, nous tenons compte dans la modélisation des possibilités d'évolution de cette corrélation selon différents régimes. L'idée principale est de partir de l'approche classique de changement de régime (cf. Hamilton [1989, 1994] et Hardy [2001]), qui concerne souvent les rendements moyens ainsi que les volatilités, pour supposer que les corrélations basculent elles aussi d'un régime à un autre.

Nous introduisons un nouveau concept qui est celui de la « corrélation à risque » correspondant à la matrice de corrélation dans le régime de crise (régime à forte volatilité des marchés). Cette matrice reflètera *a priori* la perception subjective de l'investisseur en fonction des risques auxquels il est exposé. La matrice de corrélation dans le régime à faible volatilité sera considérée comme celle reflétant la stabilité des marchés (estimée par exemple à partir de données historiques hors périodes de crises). Une illustration numérique de cette démarche est présentée.

Concernant les modèles d'allocation d'actifs, notre étude est axée sur la comparaison des modèles disponibles selon les hypothèses sous-jacentes de « rebalancement » des portefeuilles : nous faisons la distinction entre une gestion « statique » (pour laquelle les poids reviennent périodiquement à ceux de l'allocation stratégique de long-terme - cas par exemple des modèles à poids constants *Fixed-Mix*) et une gestion dite « dynamique », pour laquelle les poids peuvent s'écarter définitivement de l'allocation stratégique initiale selon des règles de gestion prédéfinies.

Nous verrons que l'approche dynamique présente l'avantage théorique de la robustesse face aux changements de régime des marchés. L'autorisation du changement des poids des différentes classes d'actifs, sur la base d'une règle de gestion bien définie, constitue *a priori* un élément intéressant. Cela en effet permet l'ajustement des expositions aux différentes classes d'actifs suite à l'évolution des conditions de marché.

La gestion dynamique de portefeuille sur le long terme reste un domaine de recherche relativement peu exploré, par comparaison avec l'importance des travaux déjà réalisés sur les aspects à court terme. Nous pouvons néanmoins citer, par exemple, les études de Hainaut et al. [2005], Hainaut et al. [2007], Rudolf et al. [2004] et Yen et al. [2003], sachant que ces différents auteurs font part des difficultés pratiques d'implémentation, en raison de la complexité des modèles proposés.

Cette réflexion sur la construction de modèles d'allocation d'actifs applicables dans une optique prévisionnelle à long-terme nous conduira à étudier une approche innovante fondée sur les techniques de « programmation stochastique » (cf. Birge et Louveaux [1997]). Il s'agit d'une version adaptée d'une technique déjà utilisée dans le domaine de l'ingénierie pour la planification de la production (cf. Dantzig et al. [1990], Escudero et al. [1993]). L'objectif sera la mise en évidence et l'étude des caractéristiques de cette approche.

Comme déjà mentionné, cette thèse accorde également une attention toute particulière aux techniques numériques de recherche de l'optimum, qui demeurent des questions essentielles pour la mise en place d'un modèle d'allocation. Le point de départ sera notre constat d'un temps de calcul significatif dû simultanément à un nombre élevé de scénarios économiques générés et à un nombre d'allocations d'actifs testées également élevé.

Dans ce cadre, nous présentons un algorithme d'optimisation globale d'une fonction non convexe et bruitée. L'algorithme est construit après une étude de critères de compromis entre, d'une part, l'exploration de la fonction objectif en de nouveaux points (correspondant à des tests sur de nouvelles allocations d'actifs) et d'autre part l'amélioration de la connaissance de celle-ci, par l'augmentation du nombre de tirages en des points déjà explorés (correspondant à la génération de scénarios économiques supplémentaires pour les allocations d'actifs déjà testées). Une application numérique illustre la conformité du comportement de cet algorithme à celui prévu théoriquement.

L'ensemble de ce travail se place dans le contexte des régimes de retraite par répartition en France : par suite les différentes applications proposées tiendront compte des spécificités de ces régimes. De même, la faisabilité opérationnelle est un des objectifs « cibles » que nous gardons à l'esprit tout au long de cette étude. Enfin, les trois années de préparation de cette thèse ont donné lieu à la rédaction de trois articles, portant respectivement sur les générateurs de scénarios économiques (cf. Faleh, Planchet et Rullière [2010] « Les générateurs de scénarios économiques : de la conception à la mesure de la qualité », Assurances et gestion des risques, n° double avril/juillet, Vol. 78 (1/2)), les techniques numériques d'optimisation (cf. Rullière, Faleh et Planchet [2010] « Un algorithme d'optimisation par exploration sélective », soumis) et l'allocation stratégique d'actifs avec les techniques de programmation stochastique (cf. Faleh [2011] «Un modèle de programmation stochastique pour l'allocation stratégique d'actifs d'un régime de retraite partiellement provisionné », soumis).

Importance de l'allocation stratégique d'actifs

L'allocation stratégique d'actifs d'un régime de retraite (ou d'une compagnie d'assurance en général) est souvent définie comme une étape d'un processus plus général de la gestion actif-passif, en particulier comme l'étape en aval de l'appréhension des risques et en amont de l'allocation tactique d'actifs. Compte tenu de ce positionnement, l'allocation stratégique d'actifs vise, soit à confirmer l'optimalité de la structure de l'actif existant de la réserve, soit à proposer une structure optimale d'actifs de cette réserve qui permette au régime d'atteindre un

certain objectif de performance financière tout en respectant ses engagements avec un niveau de confiance donné. La caractérisation de 'stratégique' vient d'une part de l'horizon temporel auquel s'appliquent les études d'allocation stratégique, d'autre part du nombre limité de classes d'actifs considérées dans ces études, généralement limité entre trois et dix au maximum.

Après avoir déterminé l'ensemble des allocations d'actifs possibles (allocations à tester), la résolution du problème de détermination de l'allocation stratégique d'actifs pour le régime de retraite étudié passe en pratique par trois étapes principales. La première étape consiste à **générer des trajectoires** pour chaque classe d'actifs (actions, obligations, immobilier, etc.). La deuxième consiste à projeter chaque allocation possible en fonction de **la règle de gestion** du dispositif. La troisième étape vise à déterminer la valeur de **la fonction objectif** pour chacune de ces allocations. Seront susceptibles d'être retenues alors celles qui à la fois maximisent la fonction objectif et respectent les contraintes fixées.

L'importance de l'allocation stratégique pour les investisseurs à long terme est mise en évidence dans l'étude de Brinson et al. [1991] comme le montre le graphique suivant :

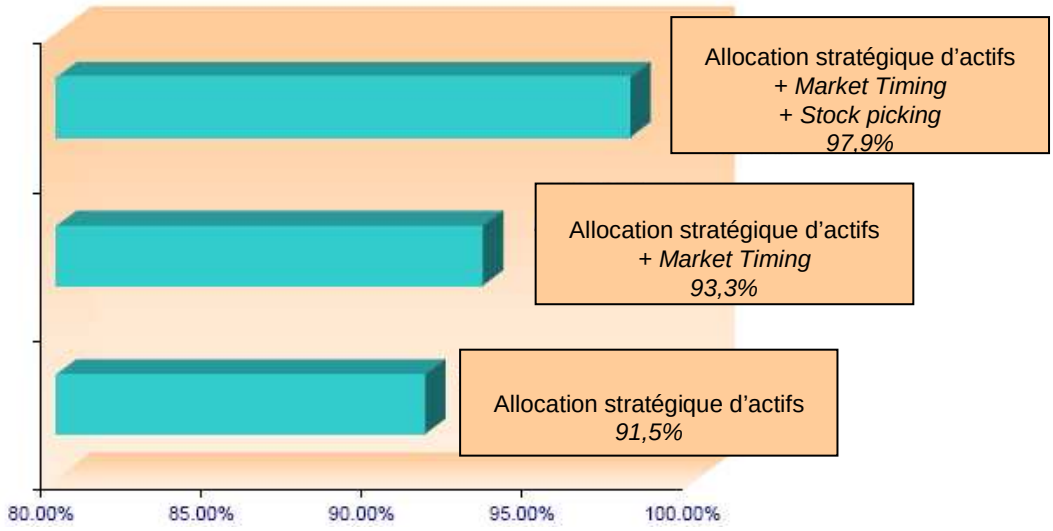


Fig. 1 : *Pourcentages de la performance globale expliqués par certaines composantes de l'allocation d'actifs selon l'étude de Brinson et al. [1991]*

Selon cette étude, la gestion passive d'un indice de référence répliquant le jeu de poids stratégiques du portefeuille suffit pour obtenir l'essentiel de la performance espérée de ce portefeuille (autrement dit l'évolution de la rentabilité d'un portefeuille est à 91,5 % le résultat de l'évolution des classes d'actifs sur lesquelles il est investi). L'allocation tactique qui permet de bénéficier des inefficiences temporaires du marché des titres et qui peut avoir pour source soit une sélection optimale de titres différents de ceux de l'indice (*stock picking*) soit le choix optimal des dates d'achat et de vente de ces titres (*market timing*), n'explique que prés de 6,5 % de la performance globale obtenue *in fine* sur le portefeuille.

L'allocation stratégique doit tenir compte de l'évolution dans le temps des rendements des actifs et de leur structure de dépendance. Cela passe par la sélection de la dynamique de chaque variable pour ensuite effectuer l'estimation des paramètres du modèle choisi (le

calibrage du modèle). Nous aboutissons ainsi à la mise en place d'un modèle de génération de scénarios économiques de long terme permettant la projection aussi bien de la valeur de l'actif que celle du passif (le cas échéant).

Allocation stratégique et générateurs de scénarios économiques

La projection sur le long terme des valeurs de marché des actifs financiers et des variables macro-économiques, souvent appelée « génération de scénarios économiques », constitue une phase cruciale dans le processus d'allocation stratégique des fonds d'une compagnie d'assurance ou d'un fonds de retraite. Elle est un élément central de l'évaluation des provisions pour les garanties financières sur des contrats d'épargne dans le cadre de la directive Solvabilité II (cf. Planchet [2009]).

Un générateur de scénarios économiques (GSE) s'avère ainsi un outil important d'aide à la décision dans le domaine de la gestion des risques, en permettant d'obtenir des projections dans le futur des valeurs des éléments qui sont présents dans les deux compartiments du bilan de la société : les actifs (actions, obligations, immobilier,..) et le passif (provisions techniques, dettes financières,..). L'obtention de ces valeurs passe par la projection d'autres variables macro-économiques et financières telles que les taux d'intérêt, l'inflation des prix, l'inflation des salaires et le taux de chômage. A titre d'exemple, les prix futurs des obligations sont déduits à partir de la projection des taux d'intérêt (cf. Ahlgrim et al. [2008]), de même les prestations futures du régime de retraite peuvent être indexées sur l'inflation des prix et/ou l'inflation des salaires (cf. Kouwenberg [2001]).

Du point de vue opérationnel, la mise en œuvre de ces projections s'est basée initialement sur des méthodes déterministes, dans des logiques de scénarios, pour tester le comportement des objectifs techniques dans différentes situations considérées comme caractéristiques. Grâce au développement concomitant de l'outil informatique et des techniques de simulation, il est devenu possible de générer à moindre coût de très nombreux scénarios en tenant compte des interactions entre les multiples sources de risque et de leurs distributions, dans une perspective probabiliste.

La construction et la mise en œuvre d'un GSE passent par les quatre étapes suivantes (voir par exemple Hibbert et al. [2001] ou Ahlgrim et al. [2005]) : la première étape consiste en l'identification des sources de risque prises en compte et des variables financières à modéliser (taux d'intérêt, inflation, rendement des actions, etc.) qu'on appellera dans la suite les variables du GSE. Ensuite est effectué le choix du modèle pour la dynamique de chacune de ces variables. La troisième étape consiste à sélectionner une structure de dépendance entre les sources de risque de façon à obtenir des projections cohérentes. Ensuite l'estimation et le calibrage des paramètres des modèles retenus doivent être effectués. Enfin l'analyse des résultats obtenus de chaque GSE se fait en termes probabilistes, en analysant la distribution d'indicateurs clés tels que le surplus, ou la valeur nette de l'actif.

Construire un GSE pertinent pour l'ensemble des problématiques techniques n'est pas chose facile, et en pratique les objectifs des décisionnaires en termes de choix de gestion ont une incidence sur la manière de structurer le générateur (cf. Ahlgrim et al. [2008]). Deux niveaux de difficultés peuvent être distingués : celui relatif à la conception théorique d'un GSE et celui relatif à sa mise en œuvre pratique. Il est donc particulièrement important de définir avec précision les principaux éléments qui caractérisent un GSE ainsi que les indicateurs de mesure de sa qualité et de sa performance.

Allocation stratégique dans le cadre de l'ALM

De son côté, la gestion actif-passif, ou *Asset Liability Management* (ALM), consiste dans une méthode globale et coordonnée permettant à une société et notamment à un régime de retraite partiellement provisionné, de gérer la composition et l'adéquation de l'ensemble de ses actifs et passifs. Les techniques utilisées pour sa mise en place diffèrent particulièrement en fonction de la nature des engagements du régime.

Récemment, la gestion actif-passif s'est imposée pour les fonds de pension comme une approche de gestion des risques qui tient compte des actifs, des engagements et aussi des différentes interactions existantes entre ces deux parties (cf. Adam[2007]). Les gérants des fonds doivent déterminer les stratégies admissibles qui garantissent avec une probabilité suffisante que la solvabilité du fonds est assurée (compte tenu des prestations attendues). La solvabilité est définie comme la capacité du fonds à payer les prestations sur le long terme.

La solvabilité du fonds à une date fixée est souvent mesurée par le ratio de financement (le *funding ratio*). Il s'agit du ratio de la valeur des actifs financiers du fonds par rapport à la valeur des engagements actualisés à un taux de rendement choisi. Le sous financement (*under-funding*) aura lieu lorsque ce ratio est inférieur à 1. Une autre façon d'exprimer le sous financement est de dire que le surplus est négatif. Ce dernier représente en fait l'écart entre la valeur des actifs et la valeur des engagements actualisés.

Notons aussi que la définition de la notion d'actif ou d'engagement diffère selon l'approche utilisée (comptable ou économique) et selon le type de régime (capitalisation, répartition provisionné ou pas) : cela influe sur le calcul du ratio de financement.

Les fonds des régimes de retraite sont plus ou moins exposés à plusieurs facteurs de risque, dont le plus important pour le gérant de fonds est le risque de sous financement (*risk of under-funding*) : c'est le risque que la valeur des engagements soit supérieure à la valeur des actifs. Les sources de ce risque peuvent être classées en deux catégories : les risques financiers (risque des actions, risque de taux, risque de crédit,…) et les risques actuariels (risque de longévité, risque de taux technique supérieur aux taux observés sur le marché,…).

En effet, le niveau du ratio de financement change au cours du temps, essentiellement à cause des fluctuations de la valeur de marché des différentes composantes du bilan. Comme conséquence, les régimes de retraite provisionnés sont, par exemple, amenés à rebalancer leur allocation d'actifs et/ou à ajuster le taux de cotisation de leurs affiliés afin de mieux contrôler le changement du niveau de ce ratio.

Dans la littérature, les modèles d'ALM sont généralement classés en trois groupes présentés chronologiquement comme suit.

Le premier groupe contient les modèles d'adossement (ou *matching*) et d'immunisation par la duration (cf. Macaulay [1938], Redington [1952]). Ces modèles se basent sur le fait que les investissements sont essentiellement effectués dans des obligations. Ceci nous permet d'obtenir, soit un adossement des flux de trésorerie des actifs financiers à ceux du passif (*matching)*, soit un adossement de la duration de l'actif à celle du passif (immunisation par la duration).

Ces techniques étaient utilisées jusqu'au milieu des années 80 et avaient comme inconvénients principaux la considération du risque de taux comme seule source de risque pour le fonds, ainsi que la nécessité d'un rebalancement périodique du portefeuille en ré-estimant à chaque fois la duration du passif, qui change continûment du fait du changement des taux d'intérêts et de l'écoulement du temps. Pierre [2009] présente une approche de la couverture de passif qui vise à résoudre les problèmes mentionnés ci-dessus. L'actif sera constitué dans ce cas d'un portefeuille de couverture de type taux (obligations ou dérivés) couplé à un portefeuille de rendement. Selon Pierre [2009], cette architecture permet une flexibilité suffisante pour permettre à l'actif de s'adapter aux mise à jour de la valeur des engagements lors de la revue des hypothèses actuarielles ayant permis leur évaluation.

Le deuxième groupe contient les modèles basés sur la simulation de scénarios déterministes et sur la notion de surplus (Kim et Santomero [1988], Sharpe et Tint [1990], Leibowitz et al. [1992]). Les modèles de surplus ont pour objet la minimisation du risque de perte du surplus (mesuré par la variance de la rentabilité du surplus) sous contraintes de rentabilité et de poids des actifs. Ils sont des modèles mono-périodiques, ce qui limite leur utilité en pratique pour des problèmes d'allocation sur le long terme.

Le troisième groupe de modèles utilise les techniques de simulation stochastique (*Monte Carlo*) pour modéliser l'évolution des différents éléments, que ce soit au niveau des actifs financiers et des engagements, ou au niveau des variables de marché et des variables démographiques (cf. Frauendorfer [2007], Munk et al. [2004], Waring [2004], Martellini [2006]). Ainsi les lois de probabilité associées aux résultats du fonds de retraite sur le long terme peuvent être estimées. A ce niveau, nous nous proposons de distinguer deux sous-groupes de modèles d'ALM basés sur les techniques stochastiques. L'élément clé de distinction sera si oui ou non les poids des différents actifs reviennent périodiquement à ceux de l'allocation stratégique définie initialement (si oui, les modèles seront appelés modèles à poids constants ou stratégie *Fixed-Mix*).

- Pour le premier sous-groupe de modèles à poids constants et malgré les avancées réalisées avec ces techniques (surtout au niveau de l'implémentation informatique), l'aspect dynamique de l'allocation stratégique reste encore marginalisé. En fait, ces modèles permettent de comparer des allocations constantes dans le temps (statiques) indépendamment des opportunités liées aux évolutions inter-temporelles des marchés (cf. Merton [1990], Dempster et al. [2003], Infanger [2002], Brinson et al. [1991] et Kouwenberg [2001]).

- Le deuxième sous-groupe de modèles, et le plus récent, est principalement inspiré de la théorie du choix de la consommation et de portefeuille développée par Merton [1971]. Il s'agit des modèles d'allocation dynamique ou inter temporels. Par exemple, à partir de la définition de la fonction objectif pour l'investisseur, ces modèles permettent la détermination d'une trajectoire des poids des différents actifs jusqu'à la date d'échéance (l'ajustement des poids est fonction des évolutions projetées du marché et de la règle de gestion prédéfinie). L'allocation stratégique retenue sera l'allocation optimale d'actifs à la date initiale t_0. Le cadre d'utilisation de ces modèles récents se heurte au problème d'implémentation vu la complexité des outils mathématiques employés (cf. Hainaut et al. [2005], Hainaut et al. [2007], Rudolf et al. [2004] et Yen et al. [2003]).

Le nécessaire positionnement de l'analyse des risques financiers par rapport au cadre Solvabilité II

Solvabilité II est une directive-cadre qui a pour but de mettre à jour le système européen de solvabilité des entreprises d'assurance. Cette directive vise en particulier une meilleure adaptation des capitaux propres exigés des compagnies d'assurances et de réassurance avec les risques que celles-ci encourent dans leur activité. Sa mise en application est prévue au 31 octobre 2012 lorsque les discussions entre la commission européenne et le parlement européen auront abouti et que cette directive aura été retranscrite dans les législations nationales par chaque parlement.

Dans cette directive, une formule standard permet le calcul du capital nécessaire pour couvrir les risques supportés par les assureurs et les réassureurs notamment suite à un choc provoqué par un évènement exceptionnel (ce niveau de capital est également appelé capital de solvabilité requis ou SCR en anglais pour *Solvency Capital Requirement*). Cela est dans le but de contrôler la probabilité de ruine à un an et de la limiter à moins de 0,5 %. Une autre alternative proposée aux assureurs sera l'utilisation d'un modèle interne complet (basé sur leur structure de risque spécifique). Dans ce cas, une validation de l'autorité de contrôle sera requise préalablement à la détermination effective du SCR à partir du modèle interne.

Le calcul du SCR tient compte d'une panoplie de facteurs de risque (taux, action, *spread*, concentration, liquidité, etc.) agrégés sous forme de modules et ce en fonction du domaine d'activité de la compagnie. Le projet de spécifications techniques pour la cinquième étude[3] quantitative d'impact (QIS5) précise que l'analyse des risques se décline en quatre catégories principales que l'on retrouve au sein des différents modules :

- le risque de marché, provenant de l'incertitude associée à la valeur et aux rendements des actifs financiers.
- le risque de souscription, provenant de l'incertitude liée à la mesure des engagements pris par l'assureur en vie, en santé et en non vie.
- le risque de contrepartie, lié à au défaut potentiel des contreparties.
- le risque opérationnel comprenant l'ensemble des risques associés aux procédures de gestion interne et aux conséquences d'un dysfonctionnement à ce niveau.

Chacune de ces grandes catégories de risques se retrouve éclatée dans les différents sous modules qu'il faut ensuite agréger, de manière à tenir compte de la dépendance entre les risques (cf. graphique 2).

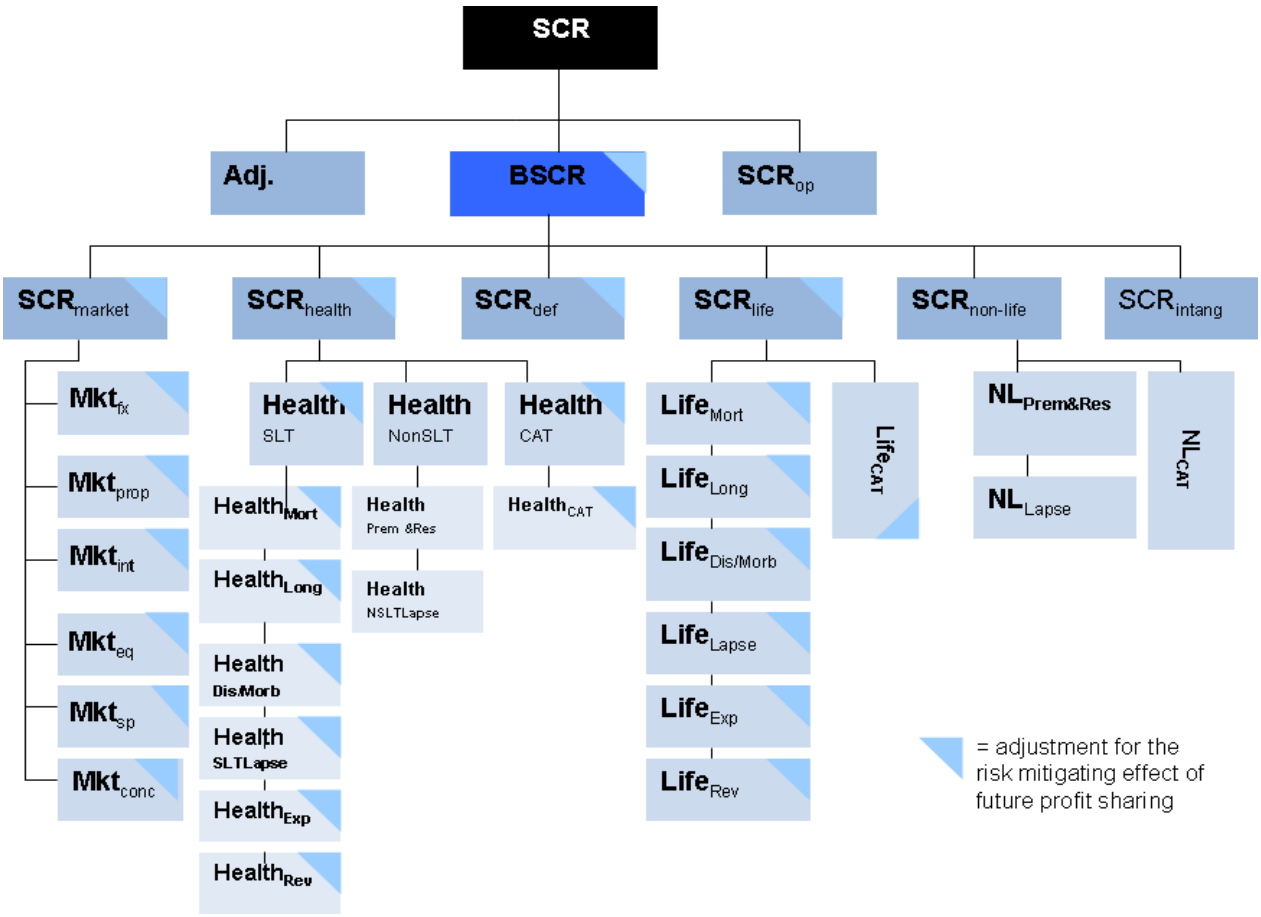
L'application de Solvabilité II aux opérations de retraite semble problématique, étant donné la particularité de ces dernières :

- Horizon de gestion : La probabilité de ruine à un an n'est pas forcément une mesure de risque pertinente pour juger de la solvabilité de l'activité de retraite, pour laquelle les engagements se font sur le long terme. Ceci laisse donc une grande marge de pilotage pour les régimes de retraite susceptibles d'amortir sur le long terme les risques à court terme.

[3] Afin de mesurer leurs implications concrètes, les directives européennes doivent faire l'objet d'une étude d'impact. C'est dans ce cadre que s'inscrit la cinquième étude d'impact (QIS 5 acronyme anglais pour *Quantitative Impact Studies* 5) pour Solvabilité II.

• Politique d'investissement : Une gestion adaptée aux contrats retraite tournée vers des actifs risqués induit un besoin en capital de solvabilité élevé. La détention d'actifs réels ou risqués, tels que les actions ou l'immobilier, est fortement pénalisée dans la formule standard, alors qu'il s'agit de supports qui présentent des avantages notables sur le long terme pour optimiser le couple rendement/risque.

Toutefois, le positionnement par rapport à la grille d'analyse des risques proposée par Solvabilité II est incontournable compte tenu de la référence que constitue ce référentiel (même si le présent travail n'est pas dans un contexte de Solvabilité II). De ce fait, l'analyse des risques supportés par un régime de retraite sera effectuée dans la suite conformément à cette grille de lecture.



Source : Projet de spécifications techniques pour la cinquième étude quantitative d'impact (QIS5)

Fig. 2 : *Les modules de la formule standard du SCR*

Présentation détaillée de la thèse

Cette thèse est composée de deux parties. La première partie est constituée de deux chapitres. La deuxième partie est constituée de trois chapitres.

La partie I traite des générateurs de scénarios économiques (GSE).

Le chapitre 1 de la partie I. Il reprend largement les points discutés dans l'article de Faleh et al. [2010].

Au niveau de ce chapitre, notre objectif est d'exposer l'état de l'art en matière d'identification des risques à intégrer à un GSE. Nous proposons ainsi différents critères de classement ainsi qu'une étude comparative des caractéristiques théoriques et des composantes de différents modèles de GSE.

L'intérêt est porté, dans un second temps, sur les aspects relatifs à la mise en œuvre d'un GSE. Nous présentons donc les caractéristiques des différentes structures schématiques de projection de scénarios utilisées en pratique. Toujours dans le cadre de la mise en œuvre, nous nous inspirons des travaux de Mitra [2006] et de Zenios [2007] pour recenser différentes méthodologies de génération des scénarios (les modèles ayant trait à la dynamique des variables du GSE).

Ensuite, nous étudions une série d'indicateurs, qualitatifs et quantitatifs, pour la mesure de la qualité d'un GSE tout en étudiant leurs limites. Nous nous référons pour cela au travail de Kaut et al. [2003]. Enfin, les principaux éléments exposés tout au long de ce chapitre sont illustrés à travers une application numérique.

Le chapitre 2 de la partie I. Un générateur de scénario adapté aux spécificités de la gestion ALM d'un régime de retraite est construit. Les éléments suivants relatifs à ce GSE sont détaillés : sa conception théorique, les modèles de diffusion retenus, les approches de calibrage de ses paramètres ainsi que sa mise en œuvre.

Notre objectif à ce niveau est d'utiliser le GSE construit en tant que moyen pour illustrer un certain nombre de résultats obtenus, tout en gardant à l'esprit les spécificités de la gestion actif-passif d'un régime de retraite. Quelques améliorations par rapport à ce qui se pratique usuellement sont apportées.

En particulier, notre étude se distingue par l'adoption d'une démarche novatrice dont l'objectif est de tenir compte de l'instabilité temporelle de la corrélation entre les rentabilités des actifs. Nous nous inspirons de l'approche classique de changement de régime (cf. Hamilton [1989], [1994] et Hardy [2001]), qui concerne souvent les rendements moyens ainsi que les volatilités, pour étendre son principe et supposer un changement de régime au niveau des matrices de corrélations aussi.

D'un autre côté, différents modèles de diffusion et de calibrage sont envisagés pour le panel des variables financières et macro-économiques qui interviennent dans ce contexte. Nous comparons les résultats d'estimation des paramètres du GSE avec ceux mentionnés dans l'étude d'Ahlgrim et al. [2005]. Ensuite, nous étudions les résultats de projection obtenus tout en proposant un certain nombre de tests pour juger de la cohérence de ces résultats.

La seconde partie est consacrée à l'élaboration de l'allocation d'actif elle-même.

Le chapitre 1 de la partie II. Il est question dans ce chapitre d'analyser les approches classiques (ou déterministes) de gestion actif-passif à savoir les modèles basés sur la notion de « duration » et de « surplus ». Nous détaillons différentes approches tout en mettant en évidence la différence entre elles. L'objectif de ce chapitre est de revoir l'état de l'art en matière de modèles d'ALM déterministe. Après avoir mis en évidence les limites de ces modèles, nous passons à l'étude de modèles plus élaborés et plus sophistiqués (objet des chapitres ultérieurs).

Le chapitre 2 de la partie II. Nous considérons ici une approche récente d'allocation stratégique d'actifs basée sur la stratégie dite « à poids constants » ou *Fixed-Mix* (cf. Kouwenberg [2001]).

Nous proposons une modélisation du régime-type de retraite et étudions certains critères d'allocation stratégique d'actifs en fonction du type du régime (provisionné, partiellement provisionné, etc.) sont étudiés. Nous passons ensuite à l'illustration de la stratégie *Fixed-Mix* avec une application sur les réserves d'un régime de retraite partiellement provisionné. Les résultats obtenus sont discutés et différents tests de sensibilité sont mises en place : ces tests sont liés principalement à l'impact des hypothèses de rendement ou de corrélation retenues.

Les difficultés rencontrées lors de la mise en place de la stratégie *Fixed-Mix*, dues essentiellement à la multiplicité du nombre de classes d'actifs considérées et au nombre de scénarios économiques simulés, nous ont mené à nous pencher sur les aspects d'optimisation numérique. Le point de départ est notre constat d'un temps de calcul significatif dû simultanément à un nombre élevé de scénarios économiques générés et à un nombre d'allocations d'actifs testées également élevé.

Dans ce cadre, une présentation détaillée des travaux menés lors de la rédaction de l'article de Rullière et al. [2010] est effectuée. Un algorithme d'optimisation globale d'une fonction non convexe et bruitée est présenté. L'algorithme est construit après une étude de critères de compromis entre, d'une part, l'exploration de la fonction objectif en de nouveaux points (correspondant à des tests sur de nouvelles allocations d'actifs) et d'autre part l'amélioration de la connaissance de celle-ci, par l'augmentation du nombre de tirages en des points déjà explorés (correspondant à la génération de scénarios économiques supplémentaires pour les allocations d'actifs déjà testées). Une application numérique illustre la conformité du comportement de cet algorithme à celui prévu théoriquement et compare les résultats obtenus avec l'algorithme de Kiefer-Wolfowitz- Blum (cf. Blum [1954], Kiefer et Wolfowitz [1952]).

Le chapitre 3 de la partie II.

Les modèles classiques d'ALM dynamiques sont explorés, notamment les techniques d'assurance de portefeuille (cf. Perold et Sharpe [1988]) et les techniques de programmation dynamique (cf. Cox et Huang [1989], Merton [1971]). A ce niveau, les techniques d'assurance de portefeuille basées sur la notion de CPPI ou *Constant Proportion Portfolio Insurance* (cf. Perold et Sharpe [1988]) sont mises en place et certains résultats relatifs à ce modèle sont étudiés. De même, les principes des techniques de programmation dynamique et leurs limites sont également mises en évidence.

Nous nous penchons par la suite sur une approche innovante fondée sur les techniques de « programmation stochastique » (cf. Birge et Louveaux [1997]). Il s'agit d'une version adaptée d'une technique déjà utilisée dans le domaine de l'ingénierie pour la planification de la production (cf. Dantzig et al. [1990], Escudero et al. [1993]). Dans ce cadre, nous mettons en place un modèle d'ALM dynamique basé sur les techniques de programmation stochastique.

Nous proposons, au cours d'une illustration numérique, une nouvelle méthodologie de génération de scénarios économiques que nous appelons méthodologie « des quantiles de référence ». Cette dernière permet de partir d'une structure linéaire de génération de scénarios (telle que décrite dans le chapitre 2 de la partie I) pour réduire la dimension du problème rencontré avec la stratégie *Fixed-Mix* tout en tenant compte de la corrélation entre les distributions des différentes variables projetées. Cela s'insère dans le cadre de la recherche d'une vision à la fois simplifiée, réelle et dynamique des stratégies possibles pour l'allocation d'actifs sur le long terme.

A travers la même application numérique, nous comparons certains résultats relatifs aux deux approches d'allocation stratégique d'actifs : celle basée sur la stratégie *Fixed-Mix* et celle basée sur les techniques de programmation stochastique. Nous testons également la sensibilité de cette deuxième approche par rapport au changement de certains de ses paramètres, toutes choses étant égales par ailleurs. Cette étude sur les techniques de programmation stochastique, dans le contexte de la gestion des réserves des régimes de retraite (en particulier ceux partiellement provisionnés), reprend largement les points évoqués dans l'article de Faleh [2011].

Tout au long de ce travail, nous mettons en évidence le lien effectif entre les recherches académiques et les besoins des systèmes de retraite en matière de gestion des risques et de respect des engagements vis-à-vis des futurs retraités.

Partie I :

LES GÉNÉRATEURS DE SCÉNARIOS ÉCONOMIQUES (GSE)

Introduction

Au départ, l'analyse des risques futurs pour les compagnies d'assurance et les fonds de retraite (ou de pension, selon la terminologie anglo-saxonne) se faisait « à la main » en essayant de répondre à la question « *et si jamais..?* » (étude de scénarios déterministes). Ceci a été suivi par des études basées sur l'adossement du passif par l'actif que ce soit au niveau des flux de trésorerie futurs ou au niveau des durations de ces deux compartiments du bilan. Grâce au développement de l'outil informatique et aux techniques de simulation de *Monte Carlo*, il est devenu possible de générer des milliers de scénarios économiques tenant compte des corrélations entre les différentes sources de risque pour ensuite analyser ces résultats en termes probabilistes.

Les travaux académiques antérieurs à 1984 traitent souvent une partie du problème rencontré par les actuaires. Les travaux sont concentrés uniquement soit sur les actions, soit sur les taux d'intérêt, soit sur l'inflation. Il y a peu de travaux qui mettent toutes ces variables dans un seul modèle en tenant compte des différentes interactions entre elles. Le travail de Wilkie [1984] constitue sans doute une référence à ce niveau. Ce professeur a présenté pour la première fois un modèle général qui inclut toutes les variables macro économiques et financières. De même, son modèle a été simple de point de vue de son implémentation, ce qui l'a rendu populaire et pendant deux décennies il a été la référence de tous les modèles postérieurs proposés.

Ce type de modèle, appelé par la suite « générateurs de scénarios économiques GSE » (en anglais *Economic Scenario Generator* ou ESG) permet par exemple de prendre en compte l'horizon long d'investissement d'un fonds de retraite provisionné et de contrôler l'influence des évolutions futures sur le choix de ses paramètres techniques (taux de cotisation, taux de prestation, etc.). Ce contrôle est souvent effectué dans le but de garantir un écart positif permanent entre la valeur des actifs financiers à une date donnée et la valeur des engagements actualisés à la même date via un taux d'intérêt de référence (l'écart est appelé le surplus). Il est noté à ce titre que les GSE constituent le cœur des modèles de gestion actif-passif appartenant à la génération des modèles d'ALM stochastiques. Le choix de l'allocation stratégique vient dans un deuxième temps refléter la fonction d'utilité de l'investisseur de long terme.

Le premier chapitre de cette partie reprend largement les points discutés dans l'article de Faleh et al. [2010]. Notre objectif est d'exposer l'état de l'art en matière d'identification des risques à intégrer à un GSE. Nous proposons ainsi différents critères de classement ainsi qu'une étude comparative des caractéristiques théoriques et des composantes de différents modèles de GSE.

L'intérêt est porté, dans un second temps, sur les aspects relatifs à la mise en œuvre d'un GSE. Nous présentons donc les caractéristiques des différentes structures schématiques de projection de scénarios utilisées en pratique. Toujours dans le cadre de la mise en œuvre, nous nous inspirons des travaux de Mitra [2006] et de Zenios [2007] pour recenser différentes méthodologies de génération des scénarios (les modèles ayant trait à la dynamique des variables du GSE).

Ensuite, nous étudions une série d'indicateurs, qualitatifs et quantitatifs, pour la mesure de la qualité d'un GSE tout en étudiant leurs limites. Nous nous référons pour cela au travail de Kaut et al. [2003]. Enfin, les principaux éléments exposés tout au long de ce chapitre sont illustrés à travers une application numérique.

Chapitre 1 : Etude des GSE

I- Présentation théorique

Cette section s'intéresse aux problèmes théoriques de modélisation liés aux GSE et présente une revue de la littérature concernant leurs différentes structures. L'objectif étant de mettre en évidence les composantes essentielles d'un générateur de scénarios économiques au niveau de sa conception théorique.

I-1 Problèmes théoriques de modélisation liés aux GSE

Comme mentionné ci-dessus, certaines problématiques liées à la conception théorique des GSE sont exposées. Ces dernières traitent de questions qui permettent l'amélioration de la performance des résultats et le développement du modèle pour l'adapter au contexte d'étude. Les problèmes de modélisation dans les GSE sont en premier lieu liés aux choix du modèle de la structure par terme des taux d'intérêt (STTI). La modélisation de la dynamique des rendements des actions vient dans un deuxième temps susciter elle aussi pas mal de questions.

I-1-1 Modèles des taux d'intérêt

En se basant sur les travaux de Roncalli [1998] et Décamps [1993], deux grandes classes de modèles de taux pour l'évaluation d'actifs financiers peuvent être présentées : les modèles d'absence d'opportunité d'arbitrage (AOA) et les modèles d'équilibre général.

I-1-1-1 *Les modèles d'absence d'opportunité d'arbitrage (AOA)*

L'évaluation dans le cadre des modèles d'AOA est de nature purement financière. Elle repose entièrement sur l'hypothèse d'absence d'opportunité d'arbitrage. L'utilisation de cette hypothèse fondamentale a permis de mettre en évidence deux approches d'évaluation par arbitrage :

- La première approche considère le prix des instruments financiers de taux comme fonction de variables d'état. Ces variables d'état –généralement le taux court pour les modèles univariés (cf. Vasicek [1977]), ou le couple (taux court, taux long) pour les modèles bivariés (cf. Brennan et Schwartz [1982]) sont supposées de dynamique exogène. Dans de tels modèles, l'hypothèse d'absence d'opportunité d'arbitrage exprime que la prime de risque du marché[4] est indépendante de la maturité du titre considéré. Le prix d'un produit obligataire est ensuite obtenu comme solution d'une équation aux dérivées partielles. La résolution probabiliste de ces équations permet une meilleure interprétation financière des formules d'évaluation.

- La deuxième approche d'évaluation par arbitrage est celle proposée par Ho et Lee [1986] et Heath, Jarrow et Merton [1992]. Le point clef de cette approche est la prise en compte de toute l'information contenue dans la structure de taux initiale en considérant comme donnée exogène la dynamique simultanée de taux ayant différentes maturités appelée aussi dynamique des taux terme contre terme. Cette

[4] Supplément de rendement exigé par les investisseurs pour avoir assumer le risque de détenir des actifs risqués plutôt que des actifs sans risque.

dynamique est choisie de façon à ce que l'hypothèse d'absence d'opportunité d'arbitrage soit respectée. Contrairement à l'arbitrage traditionnel la prime de risque du marché n'est pas spécifiée de façon exogène mais elle est définie implicitement par la dynamique des taux terme contre terme. Par ailleurs, aucune hypothèse sur la forme spécifique des prix des produits obligataires comme fonction d'une variable d'état n'est faite. Enfin, la dynamique de prix des zéro-coupons caractérise entièrement le modèle de courbe de taux.

Les modèles d'absence d'opportunité d'arbitrage sont particulièrement appropriés pour évaluer les prix des produits dérivés. Comme les dérivés sont évalués à partir des actifs sous jacents, un modèle qui capte explicitement les prix de marché de ces actifs est *a priori* plus performant qu'un modèle qui ne les prends pas en compte. Ce point a été remarqué notamment par Hull [2000] et Tuckman [2002] qui constatent que le recours à l'hypothèse AOA permet une évaluation des produits dérivés plus plausible que les approches d'équilibre.

Ho et Lee [1986] présentent un modèle en temps discret dans le cadre des hypothèses d'absence d'opportunité d'arbitrage. Ils supposent un *drift* (ou une tendance) de la variable dépendant du temps de façon à pouvoir répliquer les prix de toutes les obligations observées sur le marché.

L'équivalent du modèle de Ho-Lee en temps continu, pour le taux d'intérêt r_t , est :

$$dr_t = \theta(t)dt + \sigma dB_t$$

Avec B_t un mouvement brownien, (ce qui implique que pour deux dates t_1 et t_2 : $B(t_2) - B(t_1) \sim N(0, |t_2 - t_1|)$). Le *drift* dépendant du temps dans le modèle de Ho et Lee, $\theta(t)$, est sélectionné de façon à ce que les taux d'intérêt espérés convergent vers les anticipations données par le marché et soient reflétées dans la structure de taux observée initialement. Ce *drift* est lié principalement aux taux *forwards* implicites.

Heath, Jarrow et Morton [1992] (HJM) généralisent l'approche d'absence d'opportunité d'arbitrage : ainsi est prise en compte l'intégralité de la structure par terme et pas seulement le processus suivi par le taux court. Ils présentent alors la famille des processus de taux *forwards* $f(t,T)$ de la manière suivante :

$$df(t,T) = \mu(t,T, f(t,T))dt + \sigma(t,T, f(t,T))dB_t$$

$$Avec : f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}$$

$P(t,T)$: prix à la date t d'un zéro coupon de maturité T .
μ et σ : respectivement la tendance et la volatilité des taux *forwards*.

Le taux *forward* $f(t,T)$ est le taux, déterminé aujourd'hui, à une date future t et sur une durée future (période) $T - t$. HJM remarquent que le *drift* des taux *forwards* peut être exprimé en fonction des volatilités, ce qui conduit à faire de la volatilité le facteur prépondérant dans l'évaluation des prix des produits dérivés.

I-1-1-2 *Les modèles d'équilibre général*

L'évaluation dans le cadre d'un modèle d'équilibre général ne nécessite pas d'hypothèse sur les dynamiques de prix ou de taux. Ces dynamiques sont obtenues de manière endogène. Cette approche est celle de Cox, Ingersoll et Ross [1985] et de Campbell et al. [2005]. En effet, les modèles d'équilibre typiques se basent sur les anticipations des mouvements futurs des taux d'intérêt de court terme et non pas sur la courbe de taux observé à la date initiale. Ces mouvements peuvent être donc dérivés à partir d'hypothèses plus générales sur des variables d'état qui décrivent l'ensemble de l'économie. En utilisant un processus des taux courts, il est possible de déduire le rendement d'une obligation de long terme en déterminant la trajectoire espérée des taux courts jusqu'à la maturité de cette obligation. L'intérêt d'une telle approche est triple :

- s'assurer que les processus étudiés sont cohérents avec un équilibre général,
- analyser la déformation de la courbe des taux en fonction des chocs sur les variables économiques sous jacentes,
- fournir une spécification fondée pour l'expression des prix ou des risques de marchés utilisés dans la valorisation par arbitrage.

Un des principaux avantages du modèle d'équilibre est que les prix de différents actifs traditionnels ont des formules analytiques explicites (*closed-form analytic solutions*). Un autre avantage est que les modèles d'équilibre sont relativement faciles à utiliser. Mais, les modèles d'équilibre de structure par terme génèrent des prix de produits de taux qui sont potentiellement incohérents avec ceux observés sur le marché à la date initiale. Même si les paramètres de ces modèles peuvent être calibrés avec une précision significative, la structure par terme résultante peut générer des prix éloignés de ceux observés initialement sur le marché.

La formulation mathématique générale des modèles d'équilibre, dans le cas où il est considéré qu'un seul facteur explique ces mouvements, est la suivante :

$$dr_t = \kappa(\theta - r_t)dt + \sigma r_t^{\gamma} dB_t$$

Ce type de modèle continu est basé sur un seul facteur stochastique : le mouvement du taux d'intérêt instantané (court terme) r_t. La formule générale incorpore le phénomène dit « de retour à la moyenne ». Pour le comprendre, nous considérons le cas où le niveau actuel des taux courts r_t est supérieur au niveau d'équilibre anticipé θ. Dans ce cas, le changement du taux est prévu être négatif (baisse de r_t) de façon à converger vers θ. Dans l'autre cas, celui où le niveau de r_t est inférieur à θ, seule une augmentation des taux courts permet la convergence vers θ. Ainsi, quel que soit le niveau observé des taux courts, nous supposons une évolution de ce niveau vers θ.

La vitesse de retour à la moyenne est donnée par le paramètre κ. La deuxième partie de cette formule incorpore la volatilité inconnue des taux d'intérêt au cours du temps. Le dernier terme, dB_t, constitue la variation du mouvement brownien, de moyenne nulle et de variance dt.

L'incertitude est intégrée à travers le paramètre de volatilité σ :

- Si γ>0 : la volatilité des taux d'intérêt est liée à leur niveau.
- Si γ=0 : le modèle est équivalent au modèle de Vasicek [1977].
- Si γ=0,5 : le modèle est le processus proposé par Cox, Ingersoll, Ross [1985] appelé le modèle CIR.

Chan et al. [1992] estiment les paramètres de cette classe de taux d'intérêt et déterminent, à partir de données mensuelles allant de 1964 jusqu'à 1989, que la valeur de γ est approximativement de 1,5. Ces modèles sont appelés les modèles d'équilibre général car les investisseurs évaluent le prix de l'obligation en se référant principalement aux anticipations futures des taux d'intérêt et non à la courbe des taux observée initialement. En utilisant la trajectoire simulée du taux courts jusqu'à la maturité de ces obligations, il est possible de déterminer le rendement des obligations de long terme. Pour déterminer toute la structure par terme, l'investisseur pourra évaluer les prix des obligations de différentes maturités à partir de l'évolution anticipée des taux courts sur la durée de vie restante de l'obligation.

$$P(t,T) = E\left[exp\left(-\int_t^T r_u du \right) \right]$$

Avec $P(t,T)$ le prix d'une obligation zéro coupon à l'instant t (qui paie 1 euro à la date T).

Un des avantages les plus remarquables des modèles d'équilibre, est que les prix des obligations et les prix de certains autres dérivés de taux ont des formules analytiques explicites. Vasicek [1977] et CIR [1985] partent de la formule ci-dessus pour retrouver les prix des obligations :

$$P(t,T) = A(t,T)e^{-r_t B(t,T)}$$

Avec $A(t,T)$ et $B(t,T)$ sont des fonctions des paramètres connus κ, σ et θ. Ainsi, étant donné une réalisation de r_t, les taux exigés sur différentes maturités peuvent être obtenus.

$$R_r(t,T) = -log\{P(t,T)\}/(T-t)$$

L'inconvénient majeur des modèles d'équilibre est que la STTI (cf. page 21) résultante peut être incohérente avec les prix de marché observés, même si les paramètres sont parfaitement calibrés (cf. Ahlgrim et al. [2005]).

Hull et White [1990] utilisent le principe de *drift* dépendant du temps pour Ho et Lee [1986] pour présenter une extension du modèle d'équilibre de Vasicek [1977] et de CIR [1985]. Le modèle à un seul facteur de Hull et White est :

$$dr_t = \kappa(\theta(t) - r_t)dt + \sigma dB_t$$

Le tableau suivant illustre les points clefs de ces modèles :

	Modèles d'Absence d'Opportunité d'Arbitrage (AOA)		**Modèles d'Equilibre Général**
	Approches à prime de risque exogène	*Approches à prime de risque endogène*	*Approches entièrement endogènes*
Courbe des taux initiale	- Utilisée.	- Utilisée.	- Non utilisée.
Dynamique des taux	Exogène pour des taux de maturité spécifié (court, long,..).	Exogène dans le cas des taux terme contre terme.	Endogène : Anticipation des taux court terme futurs à partir d'hypothèses générales sur les variables qui décrivent l'ensemble de l'économie.
Prime de risque	Exogène et indépendante de la maturité.	Endogène : déduite à partir de la dynamique des taux terme contre terme.	Endogène.
Prix des produits de taux (obligations, etc.)	- Prix fonction de variables d'état (taux courts, taux longs, etc.). - Prix solution d'équations aux dérivés partielles.	- Aucune hypothèse sur la forme spécifique des prix comme fonction de variables d'état.	- Prix déterminés à partir des projections de taux courts.

Tab.1 : *Comparaison des caractéristiques des modèles de taux d'intérêts*

Une autre problématique théorique relative aux différentes variables modélisées, en particulier les taux d'intérêt, concerne le choix du nombre et de la nature des facteurs à utiliser pour la modélisation. Dans le cas particulier des taux d'intérêt, le choix définitif de ces deux éléments dépend du contexte d'application du GSE (projection de grandeurs réelles, évaluation de produits dérivés,…). La nature des facteurs peut être de sources différentes : elle peut être par exemple liée à l'horizon du taux (taux court, taux long) ou à la structure de la courbe des taux (facteur de courbure, de translation,…). A ce titre, Date et al. [2009] montrent la significativité statistique du choix de deux facteurs (par exemple les taux courts et les taux longs) pour expliquer l'évolution de la courbe de taux.

I-1-2 Modèles de rendement des actions

Concernant les actions, différentes problématiques liées à l'appréhension de la dynamique de leurs rendements montrent l'insuffisance de certains modèles adoptés, jusqu'à une période récente, par les sociétés d'assurance et les fonds de retraite. Souvent les rendements des

actions sont assumés suivre un mouvement brownien. En particulier, Black et Scholes [1973] supposent que les rendements suivent un mouvement brownien géométrique. Ceci implique que sur chaque sous période de temps, les rendements des actions sont distribués selon une loi normale et ils sont indépendants, les prix des actions suivent un processus lognormal. Dans ce cadre, si nous supposons que S_t est le prix des actions à l'instant t et que S_{t0} est le prix des actions à une date antérieure t_0, alors :

$$log\left(S_t / S_{t_0}\right) \sim N\left(\mu(t-t_0), \sigma^2(t-t_0)\right)$$

pour une moyenne μ et une volatilité σ.

Ce modèle lognormal, simple et pratique, fournit des approximations raisonnables sur des périodes courtes mais il est moins adapté aux problématiques de long terme. L'examen des données historiques relatives aux rendements des actions permet de constater, par exemple, que l'hypothèse d'une distribution normale ne permet pas de prévoir des valeurs extrêmes de rendement tel que réalisé dans le passé (cf. Ahlgrim et al. [2005], Mandelbrot [2005] et Hardy [2001]).

Un nombre important de modèles alternatifs a été proposé. Alexander [2001] recense une variété de ces modèle, incluant les processus autorégressifs généraux à volatilité conditionnellement hétéroscédastique (GARCH, cf. Bollerslev [1986]) et les analyses par composante principale (cf. Roncalli [1998]). De même, certains chercheurs proposent l'adaptation d'un modèle de changement de régime dans lequel les rendements des actions peuvent être simulés sous l'un des deux régimes suivants : un premier régime avec une hypothèse de volatilité relativement faible et une moyenne relativement élevée des rendements des actions et un deuxième régime avec une hypothèse de volatilité relativement élevée et une moyenne relativement faible de ces rendements. Une probabilité de transition entre ces deux régimes peut être déterminée *a priori* (cf. Hardy [2001]). L'hypothèse d'avoir plusieurs régimes est également envisageable. D'autres alternatives de modélisation sont proposées notamment dans le cadre des modèles discontinus (cf. Merton [1976]).

I-1-3 Autres classes d'actifs

Concernant les autres classes d'actifs, tels que l'immobilier et les produits dérivés, malgré les différents cadres d'hypothèses proposés pour leur projection, elles se heurtent souvent aux problèmes d'un historique peu profond, d'une liquidité insuffisante et de données confidentielles (cas des fonds de couverture). En matière de projection de grandeurs réelles sur le long terme, ces actifs sont souvent traités avec prudence et ne suscitent pas la grande part de l'intérêt des décideurs, en particulier dans le cas de l'allocation stratégique d'actifs d'un fonds de retraite où la priorité est souvent donnée aux actions, aux obligations et au monétaire (cf. Campbell et al. [2001]). Ceci ne remet pas en cause le potentiel que présentent ces actifs en tant que source de performance et/ou de couverture supplémentaire pour le portefeuille financier de la société ou du fonds de retraite (cf. Ahlgrim et al. [2005]).

I-2 Littérature sur les structures des GSE

La littérature sur les GSE est abondante. Nous proposons ici de classer les différents modèles en fonction de la structure de dépendance entre les variables, en distinguant deux catégories : structure par cascade et structure basée sur les corrélations. De même, pour chacun des

modèles cités, nous préciserons l'objectif qui lui était associé lors de sa conception. A ce titre, il est noté que l'utilisation d'un GSE a souvent pour finalité soit la projection sur le long terme et la prise de décision dans le cadre de la gestion des risques (dans ce cas, l'intérêt porte sur des grandeurs et des valeurs réelles) soit l'évaluation des prix d'équilibre des produits financiers sur le court terme, dit *pricing*, afin de déterminer la stratégie de marché convenable (achat, vente, etc.).

I-2-1 Modèles à structure par cascade

Une structure par cascade est définie comme une structure dans laquelle nous partons de la détermination de la valeur d'une variable (par exemple l'inflation) pour ensuite déduire les valeurs des autres variables (taux réels, rendements des actions, etc.). Jusqu'au début des années 1980, les travaux académiques traitent souvent une partie seulement du problème rencontré par les actuaires : les travaux sont concentrés sur chacune des classes d'actifs financiers (les actions, les taux d'intérêt, l'inflation,..) indépendamment des éventuelles interactions entre elles. L'intégration de ces interactions et l'étude du choix du modèle de dynamique des actifs financiers devient nécessaire pour garantir la cohérence des projections par rapport à un contexte donné.

Le travail de Wilkie [1986] marque de ce point de vue un changement majeur. Il présente pour la première fois un modèle général qui inclut toutes les variables macro économiques et financières. Ce modèle a l'avantage d'être simple à implémenter, raison pour laquelle il est rapidement devenu populaire et considéré comme la référence de tous les modèles proposés durant les deux décennies postérieures, malgré ses nombreuses limites.

La première version du modèle de Wilkie a été appliquée dans le cadre de la mesure de la solvabilité d'une société d'assurance par la *Faculty of Actuaries* [1986]. De même, un des premiers domaines d'application du modèle de Wilkie en actuariat était l'évaluation des engagements indexés sur les actions : dans ce cas nous supposons que les prestations dépendent des prix futurs des actions et que la réserve est principalement investie en obligations. De façon générale, ce modèle est plutôt cohérent avec des logiques de besoin de capital et de projection de valeur (cas de la gestion actif-passif par exemple) qu'avec des logiques de *pricing*.

Wilkie se base sur une structure par cascade, telle que décrite ci-dessus : il postule que l'inflation est la variable indépendante – « la force motrice » – du modèle dont la détermination se fait en premier lieu pour ensuite en dériver les valeurs des autres variables, principalement les dividendes, les revenus de dividende, les taux d'intérêt et la croissance des salaires. Le graphique 3 illustre le principe d'une structure par cascade. Dans ce cadre, Wilkie utilise un modèle autorégressif de premier ordre pour l'inflation. En 1995, il met à jour ce premier modèle en gardant les principes de sa structure par cascade mais en optant cette fois-ci pour une modélisation de l'inflation par un processus ARCH (*Autoregressive Conditional Heteroscedasticity*, cf. Engle [1982]). Ceci a été justifié, selon Wilkie, par la capacité de ce type de processus à tenir compte des caractéristiques des distributions historiques des données observées sur le marché de la Royaume-Uni depuis 1919.

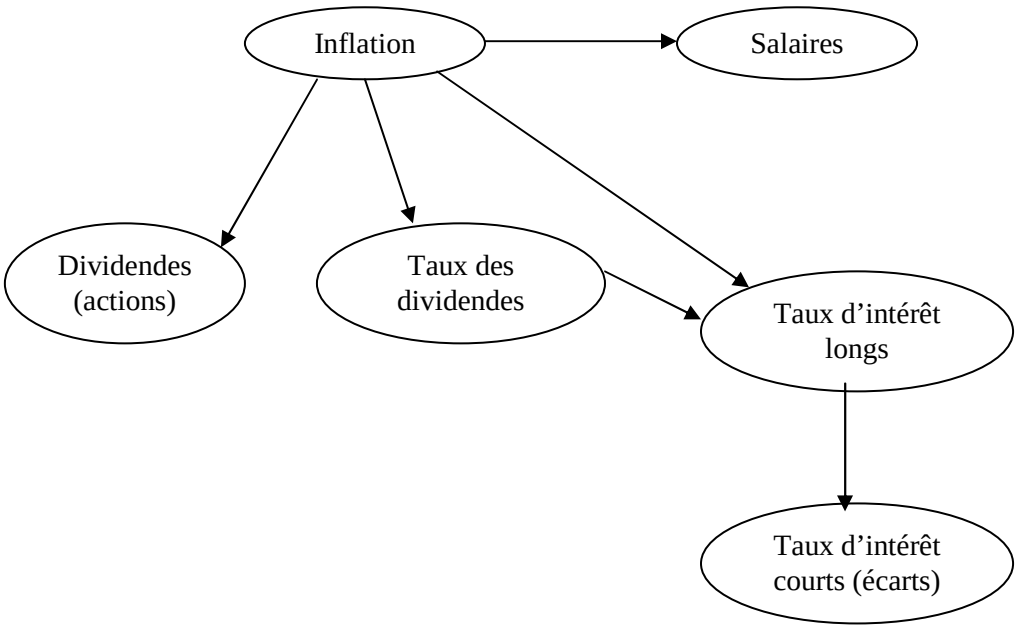


Fig.3 : *Structure par cascade dans le modèle de Wilkie [1986]*

Toutefois, l'approche de Wilkie a depuis été remise en cause, notamment du fait de sa faible capacité prédictive : il s'agit d'un modèle utilisant un grand nombre de paramètres, dont l'estimation est délicate et qui empêche de fournir des projections pertinentes. Au surplus, le modèle de Wilkie se prête mal à l'évaluation des prix des actifs dérivés, ce qui constitue un handicap important. Il est possible de se référer sur ces points à Rambaruth [2003].

Les problèmes liés au modèle de Wilkie ont également été discuté par Daykin et Hey [1990] et Huber [1995] : certains des paramètres sont instables dans le temps et une corrélation croisée significative entre les résidus des variables projetées est constatée. Le modèle des indices de prix pour l'inflation n'a pas de résidus normaux et ne permet pas la projection de périodes à chocs irréguliers avec des valeurs élevées de l'inflation. De même, la probabilité d'avoir des valeurs négatives de l'inflation avec ce modèle est élevée.

Par ailleurs, Mulvey et Thorlacius [1998] décrivent un modèle de génération de scénarios économiques appelé CAP:Link (développé commercialement par la société Towers Perrin). Ce modèle est basé sur une structure par cascade de ces variables dont la force motrice est supposée être le taux d'intérêt nominal. Il est appliqué principalement dans la gestion actif-passif de long terme. Parmi les variables clés modélisées, nous trouvons l'inflation des prix et des salaires, les taux d'intérêts de différentes maturités (réels et nominaux), le taux de rendement et les taux de dividendes des actions et les taux de change.

Les variables financières sont déterminées simultanément pour différentes économies dans un cadre d'hypothèses générales. Ce modèle s'applique ainsi aux portefeuilles de pension et d'assurance. Les dynamiques des variables sont identiques pour tous les pays alors que les paramètres sont adaptés aux spécificités de chacun d'entre eux. Les auteurs soulignent que, de façon générale, les GSE remplissent au moins l'une de ces trois fonctions suivantes : la prévision, l'évaluation (*pricing*) et l'analyse du risque. Ils considèrent que l'élément clé d'un GSE est le modèle de taux d'intérêt et supposent donc que les taux longs et les taux courts

sont corrélés à travers leurs termes de bruit blanc et que l'écart entre eux est contrôlé par un terme de stabilisation.

D'autres modèles de structure par cascade ont été développés pour l'Australie (cf. Carter [1991]), l'Afrique du Sud (cf. Thomson [1994]), le Japon (cf. Tanaka et al. [1995]) et la Finlande (cf. Ranne [1998]). L'élément commun de ces modèles réside donc dans le fait que le concepteur part de la spécification d'une structure en cascade du modèle à travers les hypothèses sur les liens de causalité entre les variables. En effet, cette structure en cascade permet un seul sens de causalité et exige du modélisateur le choix des liens les plus pertinents de point de vue économique. Par exemple, dans le modèle de Wilkie, la valeur de l'indice des prix permet de déduire la valeur de l'indice des salaires et non l'inverse. Le deuxième sens de causalité est supposé être faible (ou secondaire) sur le long terme.

I-2-2 Modèles basés sur les corrélations

La structure basée sur les corrélations repose quant à elle sur l'idée de permettre aux données disponibles (historiques) de déterminer une structure de corrélation simultanée entre les variables pour ensuite les modéliser et les calibrer en fonction de cette structure. Autrement dit, cette dernière est déterminée essentiellement à travers l'estimation des relations de dépendance observées simultanément dans le passé entre les variables modélisées (par exemple la corrélation linéaire, observée dans le passé entre les rendements des actions et l'inflation, est à retenir et à respecter lors de la projection dans le futur de ces variables). Ainsi, dans ce type de modèle, les données historiques disponibles sur les variables permettent de déduire la structure de dépendance entre elles. Les principaux modèles de GSE en littérature se sont basés sur cette structure.

En adoptant cette structure par corrélation, Campbell et al. [2001] présentent une approche dont l'application a été effectuée dans le cadre de la détermination de l'allocation stratégique d'actifs pour un investisseur de long terme, en particulier les fonds de pension. De son côté, Kouwenberg [2001] se base sur cette structure pour développer un modèle de génération de scénarios qui s'appuie sur un schéma d'arborescence pour la projection des scénarios. L'auteur compare l'effet du choix du schéma de projection sur l'allocation d'actifs optimale dans le cadre d'une gestion actif-passif d'un fonds de pension allemand. La structure d'arborescence retenue par Kouwenberg [2001] est plus adaptée à une série de modèles dynamiques de gestion actif-passif basées sur les techniques de programmation stochastique. Les caractéristiques de ce schéma de projection de scénarios sont présentées de façon détaillée à la section II de ce chapitre.

Hibbert et al. [2001] présente un autre modèle qui génère des valeurs cohérentes, selon les auteurs, de la structure par terme des taux (taux nominaux, réels et d'inflation), des rendements des actions et des revenus de dividendes. Le modèle peut être utilisé pour générer des trajectoires potentielles de chacune de ces variables dans un cadre de modélisation financière et en considérant les différentes corrélations. Hibbert et al. [2001] fournit notamment une revue intéressante des taux d'intérêt, des taux d'inflation et des rendements des actions sur les cent dernières années. Leur modèle est présenté comme un outil de planification et de prise de décisions pour les investisseurs sur le long terme et non comme un outil d'évaluation des produits dérivés (ou *de pricing*).

Ahlgrim et al. [2005] proposent enfin un modèle de GSE qui a le mérite d'être soutenu par la la *Casualty Actuarial Society* (CAS) et de la *Society Of Actuaries* (SOA), deux associations professionnelles reconnues aux Etats-Unis. Ahlgrim et al. [2005] partent essentiellement de la

critique de deux points du modèle de Wilkie [1995] : la relation entre l'inflation et les taux d'intérêt est jugée incohérente et le traitement des rendements des actions par une approche autorégressive semble trop simplificateur au regard de l'historique observé. Ils proposent des processus alternatifs en justifiant leurs choix par des *backtesting*[5] sur des données historiques profondes. Le modèle d'Ahlgrim et al. [2005] rejoint le modèle de Hibbert et al. [2001] en se présentant comme un modèle de projection de valeurs sur le long terme et de gestion des risques.

Le modèle d'Ahlgrim et al. [2005] peut être représenté par les relations suivantes :

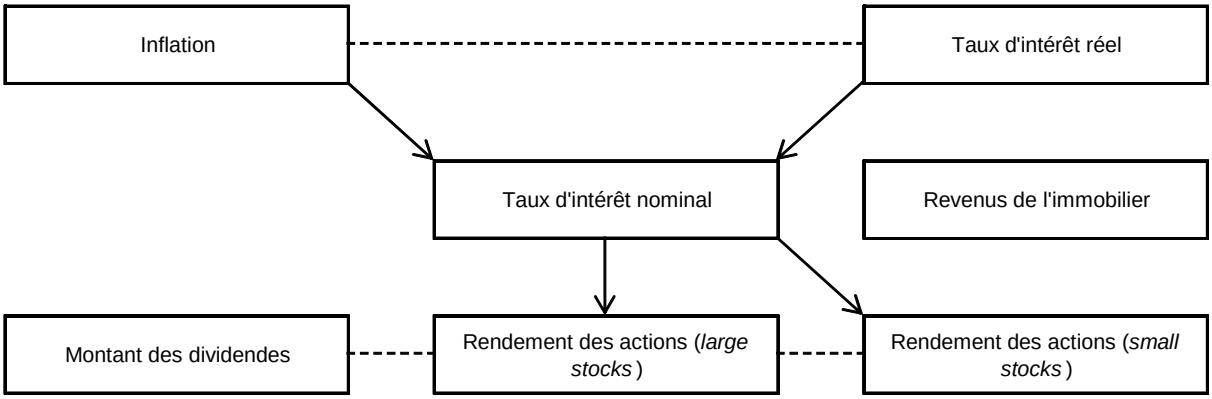


Fig. 4 : *Structure du modèle d'Ahlgrim et al. [2005]*

Ce graphique illustre le rôle prépondérant de l'inflation et du taux d'intérêt réel dans le modèle.

D'autres modèles se sont basés sur l'hypothèse que les marchés peuvent être soit faiblement efficients (c'est-à-dire les prix sur le marché reflètent toutes les informations relatives aux prix antérieurs de l'actif) soit fortement efficients (c'est-à-dire les prix sur le marché reflètent toutes les informations disponibles sur l'actif). De telles approches sont souvent appropriées pour la modélisation de court terme, en particulier pour des fins d'évaluation de produits dérivés. Pour le long terme, l'approche a moins de valeur, puisqu'elle ne tient pas compte des fondamentaux macro-économiques. Smith [1996] et Dyson et Exley [1995] présentent des modèles basés sur les principes du marché efficient pour le cas de la Grande Bretagne. Souvent, ce type de modèle cherche à exclure les opportunités d'arbitrage et ne suppose pas un retour à la moyenne pour les rendements de ces variables.

Une critique commune à tous les modèles ci-dessus, à structure par corrélation, est qu'ils sont fortement dépendants des données sur lesquelles ils sont basés. Autrement dit, si les rendements futurs relatifs à chaque variable du modèle possèdent des caractéristiques significativement différentes de celles observées sur la période historique d'estimation, le GSE pourra conduire à des projections non pertinentes. La prise en compte des avis subjectifs des experts sur le marché (sociétés de gestion, banques d'investissement, etc.) pour fixer ces niveaux futurs de dépendance constitue une source alternative d'alimentation de ces modèles.

[5] Le *Backtesting* est le test d'une stratégie sur le passé et sur un panel d'actifs financiers.

Nous le voyons, le panorama des modèles proposés dans la littérature est large. Toutefois, quelques-uns de ces travaux peuvent être synthétisés comme suit :

Structure / Objectif	Cascade	Corrélation
Projection et gestion des risques	Wilkie [1986, 1995] Mulvey et Thorlacius [1998]	Campbell et al. [2001] Kouwenberg [2001] Hibbert et al. [2001] Ahlgrim et al. [2005]
Evaluation (*pricing*)	-	Smith [1996] Dyson et Exley [1995]

Tab.2 : *Classement des principaux GSE cités dans la littérature*

Après avoir choisi la structure théorique du GSE (structure par cascade ou structure basée sur les corrélations), nous arrivons à l'étape de sa mise en œuvre pratique. Cette étape nécessite elle aussi des choix à effectuer que ce soit au niveau du calibrage des différents paramètres du GSE ou au niveau de la génération des trajectoires possibles de ses variables. Concernant le calibrage d'un GSE, différentes techniques peuvent être citées se basant sur les données historiques, les données de marché ou les avis des experts. Certaines de ces techniques seront exposées dans le chapitre 2 de cette partie (relative au modèle de GSE proposé). La section suivante s'intéresse à des problématiques liées à la génération des scénarios.

II- Mise en œuvre d'un GSE

Deux questions principales peuvent être posées lors de la mise en œuvre d'un GSE : d'une part, sous quelle forme schématique devons-nous représenter l'évolution dans le temps des scénarios futurs des variables financières et macro-économiques (inflation, rendement des actions,…) ? Et d'autre part, quelle méthodologie devons-nous adopter pour la génération de ces scénarios ? La première question concerne la structure de projection des scénarios futurs des différentes variables, tandis que la deuxième a trait au choix du modèle d'évolution des valeurs des variables du GSE (processus stochastique, *Bootstrapping*, etc.) : nous notons que ces deux éléments sont cependant liés. Cette section vise à faire l'inventaire (non exhaustif) des différentes possibilités offertes face à ces deux problématiques.

En effet, comme mentionné au début de cette partie, notre objectif sera de mettre en évidence les principaux éléments qui participent à l'amélioration de la qualité d'un GSE. Les états de sortie de ce dernier influencent directement les décisions prises en matière de gestion des risques ou d'évaluation des produits financiers. La détermination de la structure de projection des scénarios et le choix de la méthodologie de leur génération présentent deux étapes inévitables lors de la construction d'un GSE. Ils interviennent, en particulier, au niveau de la mise en œuvre opérationnelle du GSE, d'où l'intérêt de les étudier de façon détaillée et séparée.

II-1 Structure schématique de projection de scénarios pour un GSE

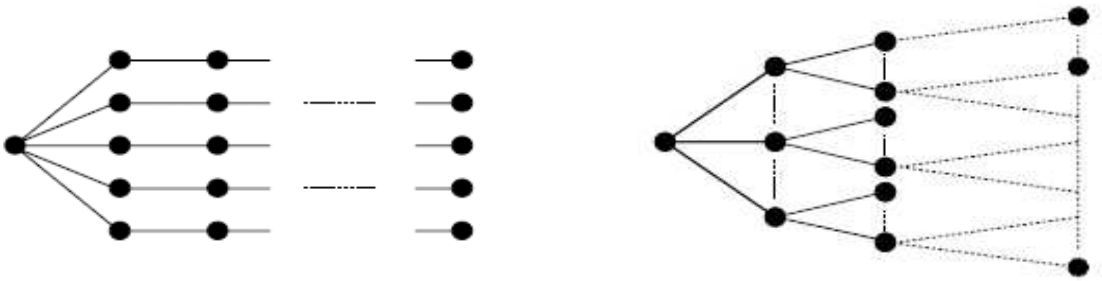
Comme mentionné ci-dessus, le choix de la structure de projection, appelée aussi structure schématique de projection, pour chacune des variables du GSE peut être considéré comme une problématique à part. Elle concerne la définition du schéma graphique de transition entre deux valeurs successives, observées à la date t et $t+1$, de la même variable. Ainsi à chaque variable du GSE peut correspondre une structure de projection particulière. Afin de simplifier l'analyse, nous supposons dans la suite que la structure de projection choisie est la même pour toutes les variables. De même, nous définissons un nœud comme la réalisation possible de la variable modélisée à une date donnée. Une trajectoire correspond ainsi à l'ensemble des nœuds successifs qui forment un scénario futur possible d'évolution de la variable financière ou macro-économique.

Dans cette section, les caractéristiques des structures de projection les plus utilisées en pratique sont présentées avec détails. En particulier, deux principales structures peuvent être avancées à ce stade : la structure de projection linéaire d'une part (cf. Ahlgrim et al. [2005]) et la structure de projection d'arbre (ou d'arborescence) d'autre part (cf. Kouwenberg [2001]). La différence principale entre ces deux structures de projection se situe au niveau de la nature de la dépendance entre les différentes trajectoires simulées. Alors que pour les structures de projection linéaire, une seule trajectoire est dérivée à partir de chaque nœud, les structures par arborescence supposent quant à elles que chaque nœud possède différents nœud-enfants et ainsi différentes trajectoires possibles sont déduites à partir de chaque nœud.

Par exemple, considérons le cas où la projection des rendements des actions se déroule sur deux périodes seulement et que pour les deux structures nous obtenons n_1 nœuds (ou rendements) à la fin de la première période. Si nous optons pour une structure linéaire de projection, il n'est possible d'obtenir que n_1 nœuds (ou rendements) à la fin de la deuxième période, chacun d'entre eux forme avec le nœud précédent une trajectoire distincte. Si par

contre nous optons pour une structure d'arbre, le nombre de scénarios à la fin de la deuxième période est m_1, avec m_1 souvent supérieur à n_1 puisque différents scénarios de rendement à partir de chacun des n_1 nœuds, simulés fin de la première période, sont projetés.

Le graphique suivant illustre la différence entre ces deux structures :



Structure linéaire des scénarios — Structure d'arbre des scénarios

Source Kouwenberg [2001]

Fig. 5 : *Comparaison entre deux structures schématiques de projection des scénarios*

La structure d'arbre des scénarios est la forme la plus récente et la plus complexe à utiliser. Notre intérêt sera focalisé dans la suite sur ce type de structure : il s'agit d'une structure plus adaptée que la structure linéaire pour l'application des techniques d'optimisation dynamique, en particulier dans le cas de la détermination de l'allocation stratégique d'actifs optimale (cf. Kouwenberg [2001]). En effet, chaque niveau dans l'arbre représente une date future (ou un moment de prise de décision) et les différents nœuds à chaque niveau représentent les réalisations possibles de la variable modélisée à cette date.

Le graphique suivant représente un exemple plus détaillé de cette structure :

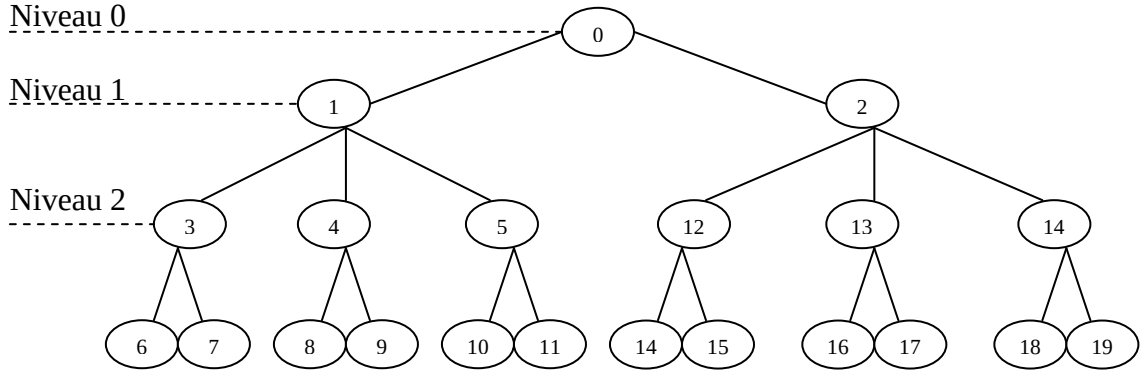


Fig. 6 : *Exemple détaillé de la structure d'arbre des scénarios*

Comme il est montré dans ce graphique, le nombre des nœuds-enfants à chaque niveau n'est pas nécessairement égal à celui du niveau suivant. Par exemple, le nœud 0 dans le graphique a deux nœuds-enfants alors que les nœuds 1 et 2 ont trois nœuds-enfants. Deux niveaux de l'arbre peuvent ne pas présenter la même période de temps. Par exemple, dans le graphique ci-dessus, le niveau 0 peut représenter le début de l'année 0, le niveau 1 la fin de la deuxième

année et le niveau 2 la fin de la dixième année. De même, dans certains arbres de scénarios complexes, tel que présenté ci-dessous, il pourrait avoir différents nombres de nœuds-enfants pour les nœuds d'un même niveau.

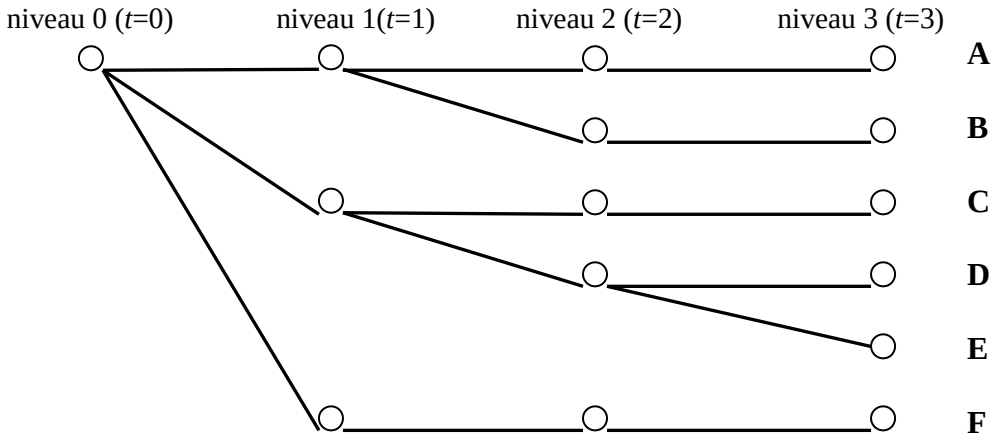


Fig.7 : *Arbre de scénarios avec différents nombres de nœuds-enfants pour les nœuds d'un même niveau*

Il existe différentes représentations mathématiques possibles de l'arbre des scénarios. Il est référé ici à la formulation de Hochreiter et al. [2002].

Considérons d'abord un processus stochastique $(\xi_t)_{[t=0,1,\ldots T]}$ discret dans le temps et continu dans l'espace et supposé représenter la dynamique des rendements des actions. L'analyse suivante peut être appliquée aux autres variables générées par le GSE. $\xi_0 = x_0$ représente la valeur d'aujourd'hui et elle est supposée être constante. La distribution de ce processus peut être le résultat d'une estimation, paramétrique ou non, basée sur des données historiques.

L'un des objectifs lors de la mise en place d'un GSE est de trouver un processus stochastique $\vec{\xi}_t$, qui prend seulement des valeurs finies et qui est aussi proche que possible du processus réel des rendements des actions : nous parlons dans ce cas de problème d'approximation. A titre d'hypothèse, le GSE est supposé avoir une structure de projection sous forme d'arbre pour ses différentes variables. Nous définissons, pour cela, l'espace d'état fini de $\vec{\xi}_t$ par S_t :

$$P\left\{\vec{\xi}_t \in S_t\right\} = 1$$

Soit $card\{S_t\}$ le cardinal de S_t. Si $x \in S_t$, nous appelons le facteur de branchement de x, le nombre des noeuds issus directement de x, c'est à dire la quantité :

$$b(x,t) = card\left\{y : P\left\{\vec{\xi}_{t+1} = y, \vec{\xi}_t = x\right\} > 0\right\}$$

Intuitivement, le processus $(\vec{\xi}_t)_{[t=0,\ldots,T]}$ peut être représenté sous forme d'arbre, avec comme racine le noeud $(x_0,0)$. Les nœuds (x,t) et $(y,t+1)$ sont connectés par un arc si $P\left\{\vec{\xi}_t = x, \vec{\xi}_{t+1} = y\right\} > 0$.

La collection de tous les facteurs de branchement $b(x,t)$ détermine la taille de l'arbre. Typiquement, le facteur de branchement est choisi avant et indépendamment de x. Dans ce cas, la structure de l'arbre est déterminée par le vecteur [$b(1)$, $b(2)$, $b(3)$,…, $b(T)$]. Par exemple un arbre [5, 3, 3, 2] a pour nombre de niveaux 5 (y compris le niveau 0) et pour nombre de nœuds 1 + 5 + 5.3 + 5.3.3 + 5.3.3.2 = 156 nœuds. Le nombre des arcs est toujours égal au nombre des nœuds moins 1. Dans ce cadre d'analyse, il est possible de considérer que la structure linéaire constitue un cas particulier de la structure d'arbre avec un facteur de branchement égal à 1 à partir de la deuxième composante du vecteur de la structure ci-dessus (c'est-à-dire [5, 1, 1, 1]).

Selon Dupacova et al. [2000], le problème d'approximation principal est un problème d'optimisation de l'un des deux types suivants et il est souvent fonction de la méthode de génération de scénarios retenue :

- Le problème à structure donnée (*The given-structure problem*) : quel processus discret $(\vec{\xi}_t)_{[t=0,\ldots,T]}$ avec une structure de branchement [$b(1)$, $b(2)$, $b(3)$,…, $b(T)$] est le plus proche d'un processus donné $(\xi_t)_{[t=0,\ldots,T]}$? Bien évidemment, la notion de proximité est à définir de manière appropriée.

- Le problème à structure libre (*The free-structure problem*) : ici aussi le processus $(\xi_t)_{[t=0,\ldots,T]}$ est à approximer par $(\vec{\xi}_t)_{[t=0,\ldots,T]}$ mais sa structure de branchement est à définir librement excepté que le nombre total des nœuds est fixé à l'avance. Ce problème d'optimisation hybride et combinatoire est plus complexe que le problème à structure donnée.

Il est à noter que la difficulté majeure lors de l'utilisation des arbres de scénarios est l'augmentation exponentielle dans le nombre de scénarios. Si trois scénarios sont générés pour chaque nœud à n'importe quel niveau parmi 21 niveaux par exemple, le nombre de scénarios générés sera 3^{20} (presque 3,5 milliards de scénarios). Le recours à l'utilisation de telle structure dans le cadre de l'allocation stratégique d'actifs fera l'objet d'une étude plus approfondie à la fin de ce travail en particulier dans le cadre de la proposition d'un nouveau modèle d'ALM (cf. chapitre 3 de la partie II).

Au delà du choix de la structure de projection pour les variables d'un GSE, l'analyse des approches possibles pour la détermination des valeurs futures des variables projetées, appelée aussi méthodologies de génération des scénarios, constitue un problème souvent rencontré par le constructeur d'un GSE. Le choix d'une méthodologie particulière n'est pas sans impact sur les résultats obtenus *in fine* (cf. Ahlgrim et al. [2008]).

II-2 Méthodologies pour la génération de scénarios économiques

Si nous nous plaçons dans le même cadre d'analyse que celui de la sous-section précédente, nous pouvons dire que cette partie vise à répondre à la question suivante: comment pouvons-nous déterminer la valeur d'un nœud. Pour cela, différentes méthodologies, ayant pour finalité la génération de scénarios économiques, peuvent être trouvées dans la littérature (cf. Kaut et Wallace [2003] et Mitra [2006]). Il est proposé dans cette sous-section de les classer en quatre groupes: les approches basées sur l'échantillonnage, les approches basées sur le *matching* des propriétés statistiques, les approches basées sur les techniques de *Bootstrapping* et les approches basées sur l'Analyse en Composantes Principales. Ce dernier groupe n'est pas indépendant des autres comme nous allons le voir dans ce qui suit.

II-2-1 Les approches basées sur l'échantillonnage

Ces approches peuvent être classées en deux sous catégories : l'échantillonnage pur univarié (ou traditionnel) et l'échantillonnage à partir de marginales et de corrélations spécifiées. Cette dernière a le mérite de générer des scénarios dans lesquels la corrélation entre les variables converge vers celle ciblée par le modélisateur.

L'échantillonnage pur est la méthode de génération de scénarios la plus connue. A chaque nœud de l'arbre de scénarios, différentes valeurs sont tirés de façon aléatoire à partir du processus stochastique $\{\xi_t\}$. Cela se fait soit par un tirage direct à partir de la distribution de $\{\xi_t\}$, soit par l'évolution du processus selon une formule explicite: $\xi_{t+1} = z(\xi_t, \varepsilon_t)$.

Dans ce cadre, la dynamique de plusieurs variables financières peut être supposée suivre un processus stochastique de type mouvement brownien géométrique, ou bien l'une de ses variantes. Les scénarios d'évolution de ces variables sont ainsi simulés à partir des hypothèses sur la discrétisation d'un processus brownien géométrique définit par exemple par:

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t)$$

$S(t)$ est le prix de l'actif, μ et σ sont respectivement le *drift* et la volatilité. Le terme $dB(t)$ est un mouvement brownien, c'est-à-dire $B(t_2) - B(t_1) \sim N(0, |t_2 - t_1|)$. Il est ainsi possible de simuler des processus stochastiques sur un intervalle de temps donné, en attribuant des valeurs aléatoires au mouvement brownien et en calculant par la suite $S(t)$.

Les méthodes traditionnelles d'échantillonnage d'une variable aléatoire permettent de constituer des échantillons seulement à partir d'une variable aléatoire univariée ; lorsque nous voulons tirer un vecteur aléatoire (correspondant à différentes variables), on aura besoin de tirer chaque composante marginale (chaque variable) de façon séparée pour les rassembler ensuite. Le résultant obtenu sera un vecteur de variables aléatoires indépendantes.

Concernant la convergence vers les moments statistiques souhaités (moyenne, variance, etc.), il existe différentes méthodes pour améliorer l'algorithme de l'échantillonnage pur. Nous pouvons par exemple utiliser les méthodes de quadrature pour l'intégration ou les suites à discrépance faible (cf. Pennanen et al. [2002]). Pour les distributions symétriques il est possible d'utiliser les échantillonnages antithétiques. Une autre méthode pour améliorer la méthode d'échantillonnage pur est d'ajuster l'arbre obtenu de façon à avoir les valeurs cibles de la moyenne et de la variance (cf. Cariño et al. [1994]).

Comme mentionné ci-dessus, les méthodes d'échantillonnage traditionnel ont des limites au niveau de la génération des vecteurs multivariés, en particulier ceux avec une corrélation spécifiée. Cependant, il existe des méthodes qui résolvent ce problème en se basant sur des approches d'échantillonnage traditionnel pour ensuite ajuster la technique de contrôle de la corrélation entre les scénarios projetés.

Ces méthodes, que nous avons appelé « échantillonnage à partir de marginales et de corrélations spécifiées » constituent donc une extension des approches d'échantillonnage traditionnel. La différence principale se situe dans le fait qu'elles permettent à l'utilisateur de spécifier à l'avance les distributions marginales ainsi que la matrice de corrélation cible. En général, il n'y a aucune restriction sur les distributions marginales, elles peuvent même appartenir à différentes familles.

A titre d'exemple, Deler et al. [2001] proposent une méthode permettant de générer des variables aléatoires à partir des séries temporelles multi-variées $\{X_t; t = 1,2,...\}$, où $X_t = (X_{1,t}, X_{2,t}, ..., X_{k,t})'$ est un vecteur aléatoire de dimension $(k \times 1)$ correspondant à l'observation à la date *t* de ces variables.

Pour cela, ils construisent un processus appelé processus de base Z_t (assimilé à un vecteur auto régressif gaussien standard) et le transforme, à travers un système de translation de Johnson [1949], en un processus ayant au moins les quatre premiers moments (moyenne, écart-type, le coefficient de dissymétrie ou *skewness* et coefficient d'aplatissement ou *kurtosis*) du processus X_t.

Il ajuste enfin la structure de corrélation de ce processus de base Z_t de façon à obtenir celle du processus X_t. L'approche utilisée dans ce cas est celle de Deler et al. [2001]. Elle se base sur la résolution d'un ensemble d'équations permettant de déterminer les corrélations à retenir entre les variables du vecteur Z_t afin de refléter la structure de corrélation cible observée entre les variables du vecteur X_t. D'autres exemples de ces méthodes se trouvent dans Dupacova et al. [2009].

Il est à noter finalement que ces approches basées sur l'échantillonnage sont utilisées dans le cas où nous avons des hypothèses sur les fonctions de distribution des composantes marginales (ou des différentes variables modélisées).

II-2-2 Les approches basées sur le *matching* des propriétés statistiques

Dans les situations où il n'y a pas d'hypothèses sur la distribution marginale du processus de génération de scénarios, les approches basées sur le *matching* des propriétés statistiques, en particulier les moments, sont les plus adaptées. Un processus de génération des scénarios par le *matching* des moments s'intéresse souvent aux trois ou aux quatre premiers moments de chacune des variables projetées (moyenne, variance, skewness, kurtosis) ainsi qu'à la matrice de corrélation.

Ces méthodes peuvent être étendues à d'autre propriétés statistiques (tel que les quantiles, etc.). Le générateur de scénarios par le *matching* des moments va ensuite construire une distribution discrète satisfaisant les propriétés statistiques sélectionnées. Par exemple, Hoyland et al. [2003] commencent par spécifier le nombre minimal de scénarios qu'il faut

générer pour ensuite obtenir l'arbre de scénarios par une optimisation non linéaire, où l'objectif est de minimiser l'erreur entre les moments théoriques de la variable et ceux fournis par l'arbre.

II-2-3 Les approches basées sur les techniques de *Bootstrapping*

Il s'agit des approches les plus simples pour générer des scénarios en utilisant seulement les données historiques disponibles sans aucune modélisation *a priori* de la dynamique d'évolution des variables du GSE (cf. Albeanu et al. [2008]). Le *Bootstrapping* se base essentiellement sur la constitution d'échantillons à partir des données observées. Dans ce cadre, les valeurs de chaque scénario représentent un échantillon de rendements d'actifs obtenu par un tirage aléatoire de certains rendements observés déjà dans le passé.

Par exemple, afin de générer des scénarios de rendement sur les dix prochaines années, un échantillon de 120 rendements mensuels tirés aléatoirement sur les 120 rendements des dix dernières années est utilisé. Ce processus est répété un certain nombre de fois afin de générer plusieurs scénarios possibles dans le futur. Autrement, dans le cas où nous avons k actifs à projeter, et en supposant qu'un historique de p périodes est disponible, nous tirons avec remise un entier entre 1 et p et nous prenons toutes les valeurs des k actifs à cette même date de façon à tenir compte de la corrélation historique entre ces derniers.

II-2-4 L'utilisation de l'Analyse en Composantes Principales

L'Analyse en Composantes Principales (ACP) est une méthode générique d'analyse de données ayant plusieurs dimensions (cf. Bouroche et al. [1980]). Elle permet l'identification des facteurs clés régissant les tendances de ces données et de réduire leur dimension tout en conservant le plus d'information possible.

Pour cela, l'ACP se base sur l'identification des vecteurs propres, des valeurs propres et des covariances. En effet, il s'agit d'une méthode descriptive qui dépend d'un modèle géométrique plutôt que d'un modèle probabiliste. L'ACP propose de réduire la dimension d'un ensemble des données (échantillon) en trouvant un nouvel ensemble de variables plus petit que l'ensemble original des variables, qui néanmoins contient la plupart de l'information de l'échantillon.

Autrement dit, pour un ensemble de données dans un espace à N dimensions, nous recherchons un sous-espace à k dimensions (défini par k variables) tel que la projection des données dans ce sous-espace minimise la perte d'information. Ces k variables seront appelés composantes principales et les axes qu'elles déterminent axes principaux. L'implémentation numérique de la méthode ACP est ainsi accessible. En pratique, l'ACP a pour objet de réduire le nombre des variables de départ du modèle pour ensuite appliquer une approche traditionnelle de génération de scénarios (parmi celles proposées ci-dessus) pour les composantes principales.

II-3 Probabilité risque-neutre Vs probabilité réelle

Les structures par terme dans les modèles d'absence d'arbitrage projettent des trajectoires de taux d'intérêt futurs qui émanent de la courbe des taux existante. L'application de tels modèles suppose que les prix sont déterminés à partir du principe d'absence d'opportunité

d'arbitrage dont découle « la probabilité risque-neutre » (cette technique est surtout intéressante lors du développement des techniques d'arbitrage).

Sous cette probabilité, le rendement de tous les actifs est supposé être le taux sans risque et la somme actualisée d'un ensemble de flux futurs est égale à la somme de ces flux actualisés au taux sans risque. Cette probabilité a pour objectif les opérations de « *pricing* » et reflète un consensus entre deux parties acceptant le transfert de risque. En fait, la méthodologie d'évaluation risque-neutre est utilisée par les banques d'investissement et les académiques pour l'évaluation des produits dérivés. Pour un modèle de retour à la moyenne, la tendance de long terme sera le taux sans risque pour toutes les variables. Nous estimons ainsi une moyenne des flux de trésorerie espérés du produit dérivé en utilisant les techniques de *Monte Carlo*. Ces flux sont actualisés au taux sans risque afin d'obtenir la valeur économique du produit dérivé.

La probabilité risque-neutre fournit des résultats cohérents par rapport à des marchés neutres face au risque mais elle n'est pas recommandée pour la mesure du risque ou pour la prévision (Adam, [2007]). Par ailleurs, les structures par terme dans les modèles d'absence d'opportunité d'arbitrage sont fréquemment plus difficiles à implémenter que leurs contreparties des modèles d'équilibre. Cette probabilité suppose un nombre d'hypothèses qui sont peu conforme au monde réel. Comme conséquence, les valeurs obtenues par cette méthodologie doivent être interprétées avec prudence. Elles peuvent constituer un benchmark intéressant, surtout lors de l'évaluation de produits dérivés par les techniques actuarielles.

La probabilité réelle a quant à elle un objectif de simulation (ou de projection) réaliste :

- Elle corrige les simulations risque-neutre et montre par exemple que sur le long terme l'investissement en obligations est plus intéressant que l'investissement en monétaire (cf. Campbell et al. [2001]).
- La simulation réelle intègre des primes de risque comme contrepartie du risque supplémentaire assumé.
- Elle corrige le calcul des indicateurs de risque (dans les scénarios de stress, analyse de sensibilité, etc.)
- Elle permet à la société de tenir compte de l'évolution de sa propre valeur de marché lors des simulations.

La probabilité réelle est utilisée lors de l'optimisation de l'exposition au risque par le gestionnaire actif-passif (Adam, [2007]). La prime de risque indique une différence entre la performance espérée de l'investisseur et le taux sans risque. Mathématiquement, le passage de la probabilité risque neutre à la probabilité réelle se fait par l'introduction d'une prime de risque μ :

$$dB_t^{réel} = dB_t^{RN} + \mu.dt$$

II-4 L'exclusion de valeurs négatives pour les modèles de taux

Beaucoup de discussions existent sur la nécessité et la façon avec laquelle nous pouvons limiter la simulation des variables négatives (Bassetto [2004]). Les variables autour desquelles se concentrent ces discussions sont les taux réels et les taux nominaux, mais l'inflation est aussi concernée. La question se complique d'avantage lorsque le taux nominal est déduit à partir des deux autres taux. Parmi les approches possibles :

- soit l'utilisation de bornes inférieures pour les taux réels et d'inflation,
- soit la fixation de bornes inférieures uniquement pour les taux d'inflation (qui peuvent donc être négatifs) : les taux réels sont quant à eux déduits dans un deuxième temps de façon à avoir des taux nominaux positifs.

III- Mesure de qualité

La mesure de la qualité d'un GSE peut avoir deux caractères : qualitatif et quantitatif. Dans les deux cas, l'objectif sera de garantir des états de sortie du modèle qui permettent de prendre les meilleures décisions.

III-1 Mesure qualitative de la qualité d'un GSE

Pour Hibbert et al. [2001], les propriétés qu'un « bon » GSE doit avoir sont les suivantes : la représentativité, la plausibilité économique, la parcimonie, la transparence et l'évolution.

Le modèle doit « imiter » le comportement des actifs financiers dans le monde réel en captant leurs principales caractéristiques (représentativité). De même, les différents scénarios générés devront être plausibles et raisonnables pour les experts du marché financier. Cela passe par l'étude de la forme de la distribution et des corrélations des différentes variables du modèle.

A ce titre, le degré de correspondance des résultats aux données historiques constitue un champ d'investigation intéressant, malgré les critiques à l'encontre de l'hypothèse de reproduction des événements historiques dans le futur. L'utilité de la comparaison des résultats obtenus par rapport aux observations déjà réalisées sur le marché peut varier en fonction des attentes du gestionnaire en termes d'évolution future du contexte macro-économique. En fait, l'objectif de similitude entre les deux séries (projetées et historiques) peut être abandonné si on juge par exemple un changement profond du contexte économique par rapport au passé (exemple : modification importante des poids des économies à l'échelle internationale, explosion de l'endettement public,…).

Le comportement joint des différentes variables simulées dans les scénarios doit aussi présenter un niveau suffisant de plausibilité et de cohérence en respectant les principes économiques. Cependant, il est à noter qu'il n'y a pas de consensus sur certaines propriétés importantes des actifs financiers. L'exemple du principe de retour à la moyenne sur le marché des actions, qui stipule que les rendements sur le long terme des actions convergent vers une tendance donnée, en est une illustration (cf. Hibbert et al. [2001]).

La parcimonie fait référence, quant à elle, à la simplicité des modèles proposés par le GSE, leur permettant d'être d'une part compréhensibles par les utilisateurs et d'autre part implémentables sur des supports informatiques. Par ailleurs, la parcimonie permet une meilleure capacité prospective.

La transparence est enfin un élément déterminant pour le succès du modèle, fortement lié à son potentiel de communication : les résultats obtenus devront être présentés sous forme de graphiques clairs et détaillés. Finalement, Le potentiel du modèle à évoluer et à élargir le champ de son étude est un élément d'attractivité supplémentaire.

Pour Zenios [2007], les critères qu'un GSE doit satisfaire afin de garantir sa bonne qualité sont au nombre de trois : l'exactitude (*correctness*), la précision (*accuracy*) et la cohérence (*consistency*).

Les scénarios devront avoir des propriétés qui sont **justifiées et soutenues par différentes publications académiques**. Par exemple, la structure par terme devra refléter le phénomène de retour à la moyenne des taux d'intérêt constaté dans plusieurs études universitaires et techniques (cf. Vasicek [1977] et Hibbert et al. [2001]). De même, la structure par terme pourra être déduite à partir des changements dans le niveau, la pente et la courbure tels que mentionné dans certaines études économiques (cf. Heath et al. [1992]).

De même, afin de vérifier la condition d'exactitude, les scénarios devront couvrir les scénarios importants observés dans le passé ainsi que tenir compte des **évènements qui ne sont pas observés avant,** mais qui ont de fortes chances d'être observés sous les conditions actuelles de marché.

Comme dans plusieurs cas, les scénarios représentent une discrétisation d'un scénario continu, l'accumulation d'un nombre d'erreurs dans la discrétisation est inévitable. Différentes approches peuvent être utilisées pour assurer que l'échantillon de scénarios représente la fonction de distribution continue sous jacente.

La précision est assurée lorsque **le premier moment jusqu'au plus grand moment de la distribution des scénarios convergent vers ceux de la distribution théorique sous-jacente**. (le *matching* des moments et des propriétés statistiques sont souvent utilisés afin d'assurer que les scénarios gardent les moments théoriques de la distribution qu'ils représentent).

La demande de précision peut mener à la génération d'un nombre élevé de scénarios. Ceci dans le but de créer une discrétisation fine de la distribution continue et afin d'accomplir la précision qu'on considère appropriée et acceptable pour le problème en question. Cependant, compte tenu des erreurs de modèle et de calibrage il est recommandé d'éviter de consacrer trop d'énergie à l'obtention d'un niveau de précision, formellement satisfaisant mais illusoire de point de vue opérationnel.

Lorsque les scénarios sont générés pour différents instruments (par exemple, les obligations, la structure par terme, etc.), il est important de voir que les scénarios sont cohérents en interne. Par exemple, les scénarios dans lesquels une augmentation dans le taux d'intérêt tient lieu simultanément avec une hausse des prix des obligations sont théoriquement non cohérents (cf. Ahlgrim et al. [2008]), même si à une échelle individuelle chacun des deux scénarios, celui du taux et celui des prix, peut être avoir lieu étant donné certains facteurs exogènes (par exemple, une augmentation des taux par la banque centrale accompagnée par l'augmentation de la demande sur les obligations comme actifs refuges sur le marché). La **corrélation** entre les différents instruments peut être utilisée afin d'assurer la cohérence des scénarios.

L'étude de la qualité de certaines méthodologies de génération de scénarios est illustrée ci-dessus :

Concernant les approches basées sur l'échantillonnage : elles permettent de probabiliser un ensemble de scénarios futurs en tenant compte de la possibilité de la reproduction des scénarios passés. Si le choix de la dynamique des variables est justifié par des références

académiques, ces approches assurent la condition d'exactitude. Par ailleurs, les conditions de précision et de cohérence peuvent aussi être préservées : avec un choix de techniques de calibrage et de discrétisation adéquates (précision) et en tenant compte de la structure de corrélation entre les variables dans les modèles (cohérence).

Nous pouvons noter que le *matching des moments* assure par définition la condition de précision puisqu'elle garde les propriétés des moments. De même, le *matching* des matrices de covariance assure la cohérence des scénarios. Cependant, cette approche est assez générale et n'est pas validée par des études académiques ce qui est nécessaire pour que l'approche respecte la condition d'exactitude.

De son coté le *bootstapping* des données historiques préserve la corrélation observée, mais ne satisfait pas la condition d'exactitude pour la génération de scénarios car il ne suggèrera pas un rendement qui n'est jamais observé dans le passé (ce qui est contre-intuitif, comme la reproduction du passé ne constitue qu'une partie des scénarios possibles dans le futur et non la totalité des scénarios potentiels, confirmation donnée par la crise économique actuelle qui a donné lieu à des rendements records jamais observés au paravent). Lorsqu'ils sont tirés de façon appropriée, les scénarios de cette approche satisfont les conditions de précision et de cohérence puisqu'ils sont en parfaite cohérence avec les observations réelles du passé.

III-2 Mesure quantitative de la qualité d'un GSE

Dans cette sous-section, inspirée des travaux de Kaut et Wallace [2003], nous essayons de présenter certains critères permettant de tester de façon quantitative la qualité d'un GSE non pas en tant qu'outil de génération de la distribution réelle de ses variables mais plutôt en tant qu'outil d'aide à la prise de décision. L'intérêt dans ce travail sera porté sur la performance des scénarios comme un *input* au modèle de prise de décision et sur la mesure de la qualité des décisions subséquentes que proposent ces scénarios (stabilité, absence de biais, etc.). L'objectif n'est donc pas la mesure du degré de correspondance des scénarios projetés par rapport à la distribution réelle des variables du GSE, une distribution qui reste peu accessible quelque soit les outils d'estimation et d'approximation employés. Nous notons aussi que la connaissance de la distribution réelle enlève le besoin de passer par un GSE pour prévoir les valeurs futures des variables financières et macro-économiques.

III-2-1 Présentation du problème

Les notations utilisées dans la suite sont identiques à celles utilisées précédemment[6]. Nous nous intéressons à la fonction objectif d'un modèle d'optimisation synthétisé comme suit :

$$\min_{x \in X} F(x, \xi_t)$$

où ξ_t correspond donc à un processus stochastique discret dans le temps et continu dans l'espace.

La résolution directe du problème réel (théorique) étant supposée être impossible, le passage par un processus discret $\{\vec{\xi}_t\}$ dont l'évolution est représentée sous forme d'un arbre ne constitue qu'une solution d'approximation (empirique). Une fois effectué, la fonction objectif devient :

[6] Le choix de la structure de projection n'influence pas l'analyse menée dans cette sous section.

$$\min_{x \in X} F(x, \vec{\xi}_t)$$

Comme mentionné ci-dessus, nous nous proposons dans ce travail de lier la performance du GSE étudié à la qualité des décisions qu'il permet d'obtenir et non pas à la performance statistique d'approximation et d'estimation du processus réel. D'autre part, il est à rappeler que l'erreur d'estimation des paramètres ne fait pas partie de l'objet de cet article ce qui ne réduit en aucun cas l'importance d'étudier la robustesse des décisions obtenues par le GSE par rapport aux choix effectués pour ces paramètres (cf. Meucci [2005]).

L'erreur, au niveau décisionnel, de l'approximation d'un processus stochastique $\{\xi_t\}$ par une discrétisation $\{\vec{\xi}_t\}$, peut être définit, soit comme la différence entre les valeurs optimales de la fonction objectif des deux problèmes : le problème réel (théorique) et le problème d'approximation (empirique), soit comme la différence entre les solutions proposées par ces deux mêmes problèmes.

Le premier type d'erreur, lié à la valeur optimale de la fonction objectif, peut être formulé comme suit :

$$\begin{aligned} e_f\left(\xi_t, \vec{\xi}_t\right) &= F\left(\arg\min_x F\left(x; \vec{\xi}_t\right); \xi_t\right) - F\left(\arg\min_x F\left(x; \xi_t\right); \xi_t\right) \\ &= F\left(\arg\min_x F\left(x; \vec{\xi}_t\right); \xi_t\right) - \min_x F\left(x; \xi_t\right) \qquad (1) \end{aligned}$$

Nous notons que $e_f\left(\xi_t, \vec{\xi}_t\right) \geq 0$, étant donné que le deuxième élément de l'écart est le vrai minimum, alors que le premier est la valeur de la fonction objectif à une solution approximée. La comparaison est effectuée non pas au niveau des solutions optimales (x^*) mais plutôt au niveau des valeurs de la fonction objectif correspondantes. En fait, notre intérêt porte en premier lieu sur la valeur minimale et non pas sur les différentes solutions qui permettent *in-fine* d'avoir cette valeur (appelée aussi la performance).

Il y a deux problèmes rencontrés dans l'équation ci-dessus de $e_f\left(\xi_t, \vec{\xi}_t\right)$:

- trouver la valeur « réelle» de $F\left(x; \vec{\xi}_t\right)$ pour une solution donnée x.
- trouver la valeur « réelle » de $F\left(x; \xi_t\right)$ qui suppose la détermination de la solution réelle x^*.

Alors que le second problème est souvent difficile à résoudre, car il nécessite la résolution du problème d'optimisation avec le processus continu réel, le premier problème peut être résolu, à travers la simulation en temps discret par exemple. Dans la sous-section suivante, nous présentons différents indicateurs de mesure de l'erreur lié à la qualité des décisions obtenues par un GSE donné. Ces indicateurs se basent sur la définition de conditions à satisfaire par la solution empirique.

III-2-2 Le test de la qualité des décisions obtenues par le GSE

Il existe au moins deux conditions nécessaires à satisfaire par un GSE afin de garantir la fiabilité des décisions déduites à partir de ses projections. La première condition est **la stabilité** de la valeur de la fonction objectif optimale par rapport aux différents arbres de scénarios : c'est-à-dire si nous générons plusieurs arbres (avec les mêmes paramètres initiaux : tendance, volatilité, etc.) et que nous résolvons le problème d'optimisation pour chacun d'eux, nous devrions avoir les mêmes valeurs optimales de la fonction objectif (mêmes performances). L'autre condition est que l'arbre de scénarios n'introduit **aucun biais** au niveau de la vraie solution $\left(x^*\right)$ et ainsi la solution proposée empiriquement doit correspondre à la solution optimale réelle. La première condition peut dans certaine mesure être testée alors qu'un test direct de la deuxième est souvent considéré comme impossible étant donné que la fonction réelle est inconnue.

III-2-2-1 *La condition de stabilité*

Deux types de stabilité peuvent être avancés : la stabilité interne de l'échantillon et la stabilité externe de l'échantillon. Ces deux types d'indicateurs interviennent au niveau de la valeur optimale de la fonction objectif et non au niveau de la solution optimale du problème.

En fait, la stabilité interne de l'échantillon signifie que si nous générons k arbres de scénarios avec la discrétisation $\{\vec{\xi}_t\}$ d'un processus donné $\{\xi_t\}$, et que si nous résolvons le problème d'optimisation pour chacun de ces arbre, nous devrions avoir (approximativement) la même valeur optimale de la fonction objectif avec comme solutions optimales x_k^*, $k = 1,...,K$.

Autrement : $F\left(x_k^*;\vec{\xi}_{tk}\right) \approx F\left(x_l^*;\vec{\xi}_{tl}\right) \qquad k,l \in 1...K$

La stabilité externe de l'échantillon concerne quant à elle les performances des arbres de la distribution réelle et elle est définit comme suit :

$$F\left(x_k^*;\xi_t\right) \approx F\left(x_l^*;\xi_t\right) \qquad k,l \in 1...K$$

Une autre formulation possible de ces deux indicateurs sera :

La stabilité interne : $\min_x F\left(x;\vec{\xi}_{tk}\right) \approx \min_x F\left(x;\vec{\xi}_{tl}\right)$

La stabilité externe : $F\left(\operatorname{argmin}_x F\left(x;\vec{\xi}_{tk}\right);\xi_t\right) \approx F\left(\operatorname{argmin}_x F\left(x;\vec{\xi}_{tl}\right);\xi_t\right)$

En utilisant l'équation (1) de la fonction d'erreur, nous déduisons que pour le cas de la stabilité externe: $e_f\left(\xi_t,\vec{\xi}_{tk}\right) = e_f\left(\xi_t,\vec{\xi}_{tl}\right)$

Il y a une différence importante entre les deux définitions : alors que pour la stabilité interne de l'échantillon nous avons besoin seulement de résoudre le problème d'optimisation basé sur les scénarios projetés par le GSE, la mesure de la stabilité externe passe par l'évaluation de la

fonction objectif réelle $F(x;\xi_t)$. En pratique, il n'est pas évident de tester précisément la stabilité externe comme nous n'avons pas la distribution réelle des variables projetées (rendements, taux, etc.).

Cela peut être démontré par l'exemple uni-périodique et uni-dimensionnel suivant[7] illustrant la différence entre les deux types d'indicateurs :

$$\min_{x\in R} F(x;\xi) = \mathrm{E}^{\xi}\left[(x-\xi)^2\right]$$

Ce problème peut être résolu théoriquement et avoir sa solution explicite pour n'importe quelle distribution de ξ :

$$\begin{aligned} F(x;\xi) &= \mathrm{E}\left[(\xi - x)^2\right] \\ &= \mathrm{E}\left[((\xi - \mathrm{E}[\xi]) + (\mathrm{E}[\xi] - x))^2\right] \\ &= \mathrm{E}\left[(\xi - \mathrm{E}[\xi])^2\right] + \mathrm{E}[2(\xi - \mathrm{E}[\xi])(\mathrm{E}[\xi] - x)] + \mathrm{E}\left[(\mathrm{E}[\xi] - x)^2\right] \\ &= Var[\xi] + 0 + (x - \mathrm{E}[\xi])^2 \end{aligned}$$

et la solution optimale est alors :

$$x^* = \underset{x\in\Re}{argmin}\, F(x;\xi) = \mathrm{E}[\xi]$$

$$F(x^*;\xi) = \min_{x\in\Re} F(x;\xi) = Var[\xi]$$

Cette solution est dite explicite parce que son obtention ne nécessite pas la simulation de la distribution de ξ et elle constitue du fait une solution générale et déterministe. Considérons le cas dans lequel nous générons des arbres de scénarios $\vec{\xi}_k$ ($k = 1...K$) et que nous obtiendrons les solutions $\mathrm{x}_k^* = \mathrm{E}\left[\vec{\xi}_k\right]$.

Nous supposons dans un premier temps que la méthode de génération de scénarios est telle que tous les échantillons $\vec{\xi}_k$ ont d'une part la même moyenne réelle $\mathrm{E}\left[\vec{\xi}_k\right] = \mathrm{E}\left[\xi\right]$ et d'autre part des variances $Var\left[\vec{\xi}_k\right]$ différentes. Ainsi $F\left(x_k^*;\vec{\xi}_k\right) = Var\left[\vec{\xi}_k\right]$ sera différente d'un échantillon à l'autre d'où une non stabilité interne. En même temps, $x_k^* = x^*$, donc $F(x_k^*;\xi) = F(x^*;\xi) = Var[\xi_k]$, ce qui confirme une stabilité externe de l'échantillon.

Si par contre, nous supposons que la méthode de génération de scénarios produit des échantillons ayant tous la même variance réelle ($Var\left[\vec{\xi}_k\right] = Var[\xi]$), mais des moyennes

[7] Pour simplifier, ξ désigne ξ_t dans cet exemple.

différentes, nous obtenons dans ce cas $F\left(x_k^*;\vec{\xi}_k\right)=Var\left[\vec{\xi}_k\right]=Var[\xi]$et la stabilité interne de l'échantillon est ainsi vérifiée. De l'autre côté, la valeur de $F\left(x_k^*;\xi\right)=Var[\xi]+\left(\mathrm{E}\left[\vec{\xi}_k\right]-\mathrm{E}[\xi]\right)^2$ diffère entre les échantillons et la résolution du problème ne vérifie pas la stabilité externe de l'échantillon.

Une question peut être posée à ce stade : quel est le type de stabilité le plus important pour juger de la performance de la méthode de génération de scénarios ? Avoir la stabilité externe sans avoir la stabilité interne signifie que la performance obtenue par approximation peut ne pas correspondre à la performance réelle. Le cas opposé, c'est-à-dire stabilité interne sans la stabilité externe, est quant à lui plus gênant pour le processus décisionnel, puisque la performance réelle des solutions empiriques dépendra de l'arbre de scénarios réel sélectionné, et nous ne pouvons pas ainsi trancher au niveau de la meilleure solution obtenue par approximation alors que dans le premier cas nous avions au moins une solution empirique stable théoriquement mais dont la performance empirique reste non stable, ce qui paraît *a priori* moins gênant.

Une autre question possible est l'étude d'une stabilité interne de l'échantillon non pas au niveau de la performance mais au niveau des solutions elles même : ainsi nous cherchons à avoir la même solution optimale quel que soit l'arbre de scénarios généré.

Dans l'exemple ci-dessus, nous pouvons remarquer qu'il est possible d'avoir une non stabilité interne de la valeur optimale de la fonction objectif (différentes variances empiriques $Var\left[\vec{\xi}_k\right]$) tout en gardant une stabilité interne au niveau des solutions elles mêmes (les solutions x_k^* = $\mathrm{E}[\xi]$ sont les mêmes pour tous les arbres). Ceci permet de déduire immédiatement la stabilité externe de l'échantillon. Ainsi, si nous détectons une stabilité interne de l'échantillon au niveau de la fonction objectif, nous devrions regarder aussi au niveau des solutions. Par contre l'autre sens d'analyse n'est pas toujours vrai et nous pouvons avoir une stabilité externe même si la stabilité interne au niveau des solutions n'est pas vérifiée.

Nous pouvons conclure *in-fine* que la stabilité est une condition minimale à respecter par la méthode de génération de scénarios. Ainsi, avant de commencer à travailler avec un nouveau modèle d'optimisation, ou une nouvelle méthode de génération de scénarios, il sera judicieux d'effectuer les tests de stabilité : le test de la stabilité interne de l'échantillon et dans la mesure du possible le test de la stabilité externe de l'échantillon.

III-2-2-2 *La condition d'absence de biais*

En plus de la condition de stabilité (à la fois interne et externe de l'échantillon), la méthode de génération de scénarios ne devra pas introduire de biais au niveau de la solution proposée. Autrement, la solution du problème d'approximation basé sur les scénarios fournis par notre GSE :

$$\vec{x}^* = \underset{x}{argmin}\, F\left(x;\vec{\xi}_t\right)$$

devra être une solution optimale du problème original. Ainsi, la valeur de la vraie fonction objectif pour la solution obtenue, $F\left(\bar{x}^*;\bar{\xi}_t\right)$, devra être approximativement égale à la valeur optimale réelle $\min_x F(x;\xi_t)$:

$$F\left(\bar{x}^*;\xi_t\right) = F\left(\arg\min_x F\left(x;\bar{\xi}_t\right);\xi_t\right) \approx \min_x F\left(x;\xi_t\right)$$

En utilisant l'équation (1) de la fonction d'erreur, cette condition s'exprime comme suit :

$$e_f\left(\xi_t,\bar{\xi}_t\right) \approx 0 .$$

Le problème est que le test de cette propriété n'est pas possible dans la plupart des cas, étant donné que cela nécessite la résolution du problème d'optimisation avec le processus continu réel, supposé jusque là comme inconnu.

III-3 Application numérique

Afin d'illustrer les points développés jusque là, nous présentons dans ce qui suit un exemple numérique simplifié prenant le cas d'un investisseur (compagnie d'assurance par exemple) intéressé par la répartition de son capital entre *n*=3 actifs financiers sur le marché européen : les actions, les obligations et le monétaire. Chacun de ces trois classes d'actifs sera représenté par un indice dont la série de données historiques mensuelles, entre le 01/01/1999 et le 31/10/2009, est récupérée à partir de la base de données Bloomberg : ainsi nous avons retenu l'indice EONIA pour le monétaire, l'indice obligataire EFFAS-Euro pour les obligations et l'indice DJ Eurostoxx50 pour les actions. Nous supposons aussi que ce problème d'optimisation de portefeuille est mono-périodique : la projection des valeurs de rendement de chacun des trois actifs se fait sur une seule période (par exemple une année).

Au niveau du choix de **la méthodologie de génération des scénarios** du GSE, un processus stochastique reflétant une hypothèse de normalité des rendements est associé à chacun des trois indices, soit :

$$R^i = \mu_i + \sigma_i \varepsilon_{t,i} \qquad \text{avec } i = 1,2,3$$

R^i correspond au rendement de l'indice i. La tendance (ou la moyenne annualisée) μ_i et la volatilité σ_i des rendements des trois indices sont estimées sur une base historique (elles correspondent respectivement à la moyenne annualisée et à la volatilité empiriques de chacune des trois séries de données retenues). $\varepsilon_{t,i}$ suit une loi normale de moyenne nulle et de variance égale à 1.

La structure de dépendance entre les variables du GSE considéré est bien une structure par corrélation. Les trois indices sont simultanément dépendants l'un de l'autre lors de leur projection et le niveau des corrélations est déterminé à travers les corrélations fixées en amont entre leurs bruits $\varepsilon_{t,i}$ $(i = 1,2,3)$.

La fonction objectif de notre investisseur sera définit comme suit :

$$Min \quad F(R_p;\varepsilon_t) = \sigma(R_p)$$
$$tel\ que \quad p \in U$$
$$E(R_p) \geq m_0$$

Avec R_p le rendement total du portefeuille composé des trois actifs (monétaire, obligations et actions), $E(R_p)$ la valeur espérée du rendement du portefeuille, $\sigma(R_p)$ la volatilité du rendement du portefeuille et m_0 la valeur minimale du rendement espéré du portefeuille (égale à 4 % dans notre cas). En pratique, l'espérance sera estimée par la moyenne empirique.

Rappelons que pour déduire le rendement du portefeuille sur une période donnée, la formule suivante est utilisée :

$$R_p = \sum_{j=1}^{n=3} p_j R^j$$

Avec p_j et R^j correspondants respectivement au poids et au rendement de l'actif *j* sur la même période.

U correspond à l'univers des allocations (ou compositions de portefeuille) à tester. La détermination de cet univers passe par la fixation d'un pas, par exemple de 5 %, pour les poids des différents actifs. Autrement :

$$U = \left\{ (p_1,...,p_n) \setminus \sum_{j=1}^{n} p_j = 1{,}1 \geq p_j = c \times 0{,}05 \geq 0, \forall j \in \{1,...,n\}, c \in \{0{,}1,...,20\} \right\}$$

Dans notre cas, 3 actifs et un pas de 5 %, nous obtenons 231 allocations à tester.
$F(R_p;\varepsilon_t)$ correspond à la fonction objectif réelle présentée dans la sous-section précédente comme $F(x;\xi_t)$[8].

La structure de projection adoptée dans cette application est une structure linéaire et ce afin de simplifier la méthode de résolution de la fonction objectif. Cette structure peut être illustrée par exemple par le graphique 8 correspondant à la variable rendement des actions projetée sur une période de 20 ans avec 500 trajectoires. Dans ce graphique, $R_{i,t}^{actions}$ représente le rendement de l'indice des actions entre la date *t* et *t-1* à la $i^{ème}$ trajectoire (ou scénario).

La taille de la structure de projection sera définie comme le nombre de trajectoires simulées dans cette structure. Un ensemble de scénarios désignera l'ensemble des trajectoires obtenues d'une variable dans le cas d'une structure de projection linéaire (équivalent de l'arbre de scénarios dans le cas de structure par arborescence).

Dans notre étude nous envisageons de procéder comme suit : nous fixons différentes tailles des structures de projection adoptées dans ce GSE : par exemple 50, 1000, 5000 et 10 000 trajectoires. Pour chacune de ces valeurs, nous faisons tourner ce GSE *m*=30 fois en

[8] La caractérisation « réelle » vient du fait que dans notre cas la résolution du problème passe par la simulation et la projection de scénarios stochastiques pour les différents actifs: on parle dans ce cas d'une résolution empirique basée sur la méthode *Monte Carlo* et non d'une résolution réelle déterminée généralement à l'aide de formules explicites.

résolvant à chaque fois le problème d'optimisation présenté ci-dessus. 120 valeurs optimales empiriques de la fonction objectif sont ainsi obtenues (30 valeurs pour chacune des 4 tailles proposées). Les valeurs de ces différents paramètres (taille, nombre de fois qu'on fait tourner notre GSE) sont ici choisies de façon arbitraire. L'idée est de tester la stabilité et la distribution de la valeur optimale empirique de la fonction objectif ainsi que de la solution empirique optimale (c'est-à-dire l'allocation optimale d'actifs) par rapport à la taille de la structure. Cette étude essayera de juger *in fine* la qualité des solutions fournies par notre GSE pour le preneur de décision.

Avant d'exposer et d'analyser les résultats obtenus, il est à rappeler que la stabilité interne de l'échantillon nécessite que, pour une taille de structure donnée, les valeurs optimales de la fonction objectif ne dépendent pas de l'ensemble de scénarios projetés. L'idéal sera donc, dans ce cas, d'avoir, pour une taille de structure donnée, 30 valeurs optimales approximativement identiques de la fonction objectif. Pour la stabilité externe de l'échantillon, les valeurs réelles de la fonction objectif (ici la variance) doivent être approximativement les mêmes pour les différents ensembles de scénarios projetés.

De même, ces valeurs devront être approximativement égales à celles obtenues dans le test de stabilité interne. Comme mentionné précédemment, cette stabilité externe de l'échantillon peut être déduite à partir de la stabilité interne de l'échantillon au niveau de la solution empirique retrouvée (l'allocation optimale dans notre cas).

Dans le cadre du test de la stabilité interne de l'échantillon, le tableau 3 présente les statistiques relatives aux 30 valeurs optimales de la fonction objectif et ce pour les différentes tailles proposées.

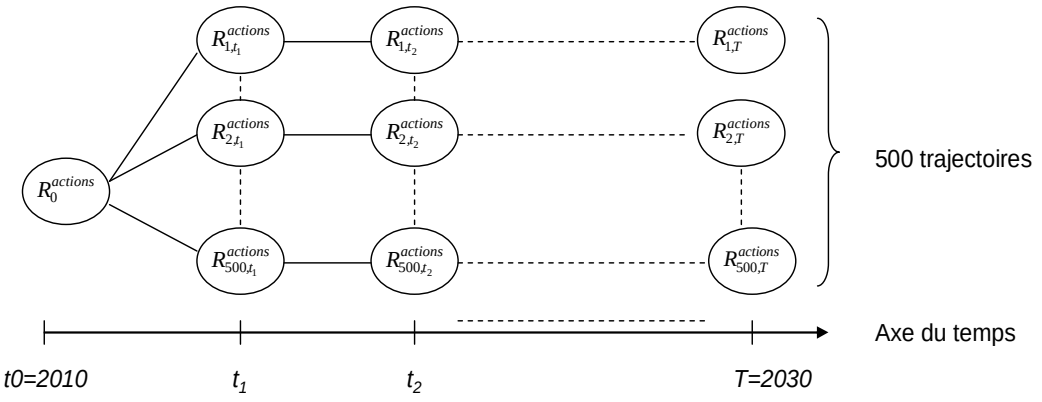


Fig. 8 : *Projection du rendement des actions selon une structure schématique linéaire*

Description du test			Nombre de scénarios			
Type de test	Fonction objectif	Valeur	**50**	**1000**	**5000**	**10 000**
Interne	$F\left(R_p ; \bar{\varepsilon}_k\right)$	Moyenne Ecart-type	0,0288 0,0050	0,0275 0,0011	0,0275 0,0007	0,0277 0,0005

Tab. 3 : *Propriétés statistiques des valeurs de la fonction objectif pour les différentes tailles retenues des structures de projection de scénarios*

De même, le graphique ci-dessous illustre la distribution de ces valeurs optimales (médiane, quantile 25 % et 75 % ainsi que le minimum et le maximum sur les 30 arbres de chaque taille).

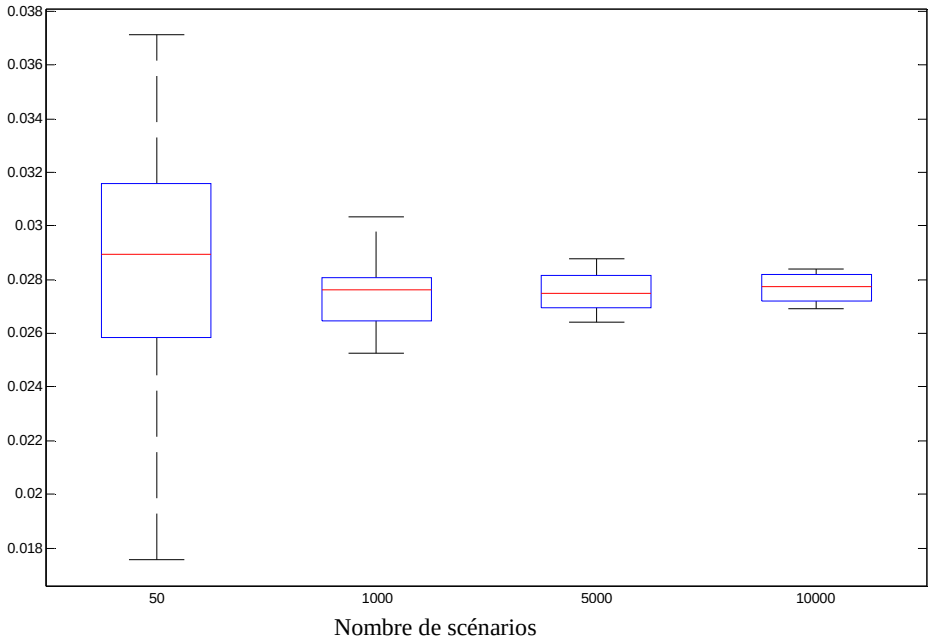


Fig. 9 : *Box plot de la distribution des valeurs de la fonction objectif pour différentes tailles de scénarios*

Comme nous pouvons le constater, plus le nombre de scénarios augmente plus la valeur optimale de la fonction objectif a tendance à baisser et à se stabiliser autour d'une valeur centrale (en particulier à partir de 5 000 scénarios). La condition de stabilité interne de l'échantillon peut ainsi être respectée par notre GSE à partir d'une certaine taille de la structure.

D'un autre côté, si la variation du poids optimal de chaque actif est relativement faible (bornée entre deux valeurs proches) cela garantit la stabilité externe de l'échantillon. Dans ce cadre, nous analysons les 30 vecteurs d'allocations optimales obtenus pour les ensembles de scénarios ayant pour taille 10 000. Les statistiques de ces solutions sont présentées dans le tableau suivant, ainsi que le graphique correspondant ci-dessous, et permettent de constater relativement une stabilité externe de l'échantillon (vu que les écarts types et/ou les étendus des poids des différents indices sont relativement faibles).

	Moyenne	Ecart-type	Etendue	Minimum	Maximum
Monétaire	0,37	0,0252	0,05	0,35	0,4
Obligations	0,55	0,0497	0,1	0,5	0,6
Actions	0,08	0,0254	0,05	0,05	0,1

Tab. 4: *Propriétés statistiques des poids des 30 allocations optimales obtenues pour la taille de 10 000 scénarios*

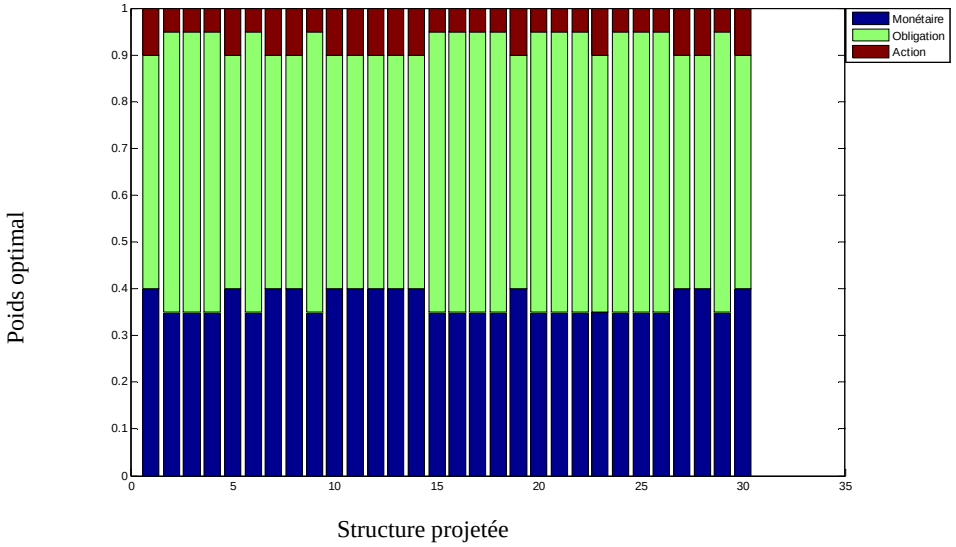


Fig. 10 : *Les allocations optimales obtenues pour les différents ensembles de scénarios (30 ensembles) ayant chacun la taille de 10 000 scénarios*

Conclusion du chapitre 1

Les résultats obtenus à travers cette application constituent des indicateurs que le décideur peut utiliser pour justifier la stabilité de sa décision finale par rapport à son générateur de scénarios : les graphiques et les tableaux présentés confirment la possibilité d'avoir une stabilité interne et une stabilité externe de l'échantillon. En pratique, les hypothèses de travail et les variables de risque prises en compte dans le processus décisionnel sont évidemment beaucoup plus nombreuses et plus compliquées mais cela ne limite pas l'intérêt de ces indicateurs comme des outils d'aide à la mesure de la performance de la décision.

La prochaine étape est une illustration plus globale du processus de construction d'un GSE. Le contexte de cette construction est celui de la projection sur le long terme de grandeurs réelles pour des fins de gestion actif-passif (cas des fonds de retraite et de certains contrats d'assurance). Le but du chapitre suivant est donc d'exposer la démarche suivie pour répondre à ce besoin ainsi que les points qui, en pratique, nécessitent une attention particulière de la part du modélisateur.

Chapitre 2 : Construction d'un générateur de scénarios économiques

I- Conception et composantes

Le générateur de scénarios économiques construit dans l'étude suivante tient compte des relations entre les différentes variables économiques et financières. La méthodologie de mise en place du modèle ainsi que les techniques de calibrage utilisées sont également exposées.

L'objet de l'utilisation de ce GSE sera la projection sur le long terme de différentes variables financières. Il s'agit en quelque sorte de préparer le terrain au fonctionnement des modèles de gestion actif-passif d'un régime de retraite comme nous allons le voir dans la partie II.

Nous supposons que les actifs constituant l'univers d'investissement sont les suivants : le monétaire, les obligations d'Etat à taux fixe (de différents niveaux de duration), les obligations d'Etat à taux indexé à l'inflation (de différents niveaux de duration), les obligations non gouvernementales à taux fixe (de duration moyenne 5 ans), les actions et l'immobilier.

Univers d'investissement	
Monétaire	-
ZC 5 ans (nominal)	obligations d'Etat à taux fixe de maturité 5 ans
ZC 10 ans (nominal)	obligations d'Etat à taux fixe de maturité 10 ans
ZC 15 ans (nominal)	obligations d'Etat à taux fixe de maturité 15 ans
ZC 8 ans (réel)	obligations d'Etat à taux indexé à l'inflation de maturité 8 ans
ZC 15 ans (réel)	obligations d'Etat à taux indexé à l'inflation de maturité 15 ans
Crédit 5 ans	obligations non gouvernementales à taux fixe de maturité 5 ans
Actions	-
Immobilier	-

Tab.5 : Univers d'investissement considéré dans l'étude

I-1 Conception théorique

Une première étape dans la conception théorique du GSE objet de l'illustration vise à arbitrer entre un modèle par cascade et un modèle basé sur les corrélations.

Comme il est mentionné dans la sous-section I-2 du chapitre précédent, les modèles basés sur les corrélations s'appuient sur des descriptions *ad hoc* de chaque classe d'actif et les agrègent ensuite pour proposer une description globale de l'actif. Dans le cadre d'un modèle basé sur les corrélations, chaque classe est ainsi modélisée finement ce qui, dans le cadre d'une démarche de gestion actif-passif (ALM), permet de décrire de manière précise l'allocation effectuée.

Les modèles par cascade proposent une description structurée de plusieurs classes d'actifs à partir en général d'une variable explicative de référence (comme par exemple le modèle de Wilkie [1995] qui s'appuie sur une modélisation de l'inflation dont découlent les taux et les

prix des actions). Les classes d'actifs tels que les taux d'intérêt peuvent apparaître modélisées de manière relativement sommaire en comparaison à la précision des modèles basés sur les corrélations.

Toutes ces observations, ainsi que celles présentées dans la sous-section II-2 du chapitre précédent, nous ont conduits à retenir un modèle proche de celui d'Ahlgrim et al. [2005] plutôt qu'un modèle du type Wilkie [1995] dans le cadre de la présente étude. Le modèle GSE développé ne supposera pas l'existence d'une structure par cascade entre ces variables.

Concernant le choix d'un modèle de structure par terme des taux, ce dernier doit tenir compte de l'objectif de planification stratégique de long terme qui diffère d'un objectif d'application de court terme (basé sur la comparaison des prix obtenus avec ceux des produits dérivés observés sur le marché). Dans cette étude, nous n'envisageons pas d'utiliser le modèle de GSE, objet de l'illustration, pour des opérations de « *trading* ». Nous cherchons plutôt à fournir aux fonds de retraite un ensemble de scénarios de taux d'intérêt et de rendement d'actifs qui soit plausible par rapport aux tendances futures et aux réalisations passées. Un cadre de modélisation avec des hypothèses **d'équilibre général** sera favorisé. Les différents actifs n'auront pas le même rendement espéré et une prime de risque sera accordée à chacun d'entre eux. En choisissant la structure par terme des taux, nous essayeront également de respecter les points mentionnés ci-dessous relatives à l'étude historique des mouvements de la courbe des taux.

L'étude historique des mouvements de la courbe des taux met en relief les points suivants (Martellini et al. [2003]) :

- les taux d'intérêt nominaux ne sont pas négatifs.
- les taux d'intérêt sont affectés par des effets de retour à la moyenne.
- les taux longs et les taux courts ne sont pas parfaitement corrélés.
- les taux à court terme sont plus volatiles que les taux à long terme.
- les courbes de taux d'intérêt peuvent avoir différentes formes (quasi-plate, croissante, décroissante, avec une bosse, avec un creux).

Dans le modèle développé, la structure par terme des taux nominaux est déduite à partir de la structure par terme des taux réels et de la structure par terme des taux d'inflation anticipés. Nous avons choisi de modéliser ces deux dernières par le modèle de Hull et White à 2 facteurs [1994] appliqué dans un cadre d'équilibre général, avec un principe de retour à la moyenne et avec un *drift* dépendant du temps. Il se base sur le principe de retour à la moyenne des taux d'intérêt. Ce modèle a le mérite d'être utilisé pour des fins de projection de long terme et dans un cadre d'analyse d'un « monde réel » (Ahlgrim et al. [2005]). Il fournit aussi des formules explicites pour les prix des obligations zéro-coupon ce qui permet à son utilisateur d'obtenir les taux de rendements de différentes maturités, à n'importe quelle date durant la période de simulation (courbe des taux).

Le choix d'un modèle multi-factoriel garantit une flexibilité au niveau de la forme de la courbe des taux. Il est montré d'un autre côté que le pouvoir explicatif d'un modèle à deux paramètres est suffisant pour la prévision des changements de taux d'intérêt (Date et al. [2009]). Ces deux facteurs retenus seront le taux long (sur un horizon long, 10 ans par exemple) et le taux court ou le taux instantané (sur un horizon court, 3 mois par exemple). Nous pouvons ainsi obtenir différentes formes de la courbe des taux en fonction de l'évolution de sa partie de long terme par rapport à sa position initiale. De même, en considérant deux facteurs, nous évitons la limite du modèle à un seul facteur qui est celle

d'une corrélation parfaite, égale à 1, entre les taux de rendement exigés sur différentes maturités. Ainsi, les taux courts et les taux longs n'auront pas à évoluer toujours dans un « *lock-step* ».

En ce qui concerne les actions, différents auteurs ont soulevé la problématique des régimes qui caractérisent le processus d'évolution de la rentabilité des actifs financiers. Le GSE que nous allons mettre en place ne suppose pas le retour à la moyenne pour les rendements des actions. La dynamique des rendements des actions se base sur le modèle de changement de régime tel que présenté par Hardy [2001]. Les actions peuvent évoluer dans le cadre de l'un de deux régimes suivants : un régime à espérance de rendement élevée et à volatilité faible des rendements des actions et un régime à espérance de rendement relativement faible (voire même négative) et à volatilité relativement élevée. Ce modèle permet de générer un *skewness* (coefficient d'asymétrie) et un *kurtosis* (coefficient d'aplatissement) plus proche de ceux de la distribution empirique des rendements des actions (Hibbert et al. [2001]). Le rendement des actions sous chaque régime est supposé suivre un mouvement brownien géométrique tel que celui de Black et Scholes [1973].

L'un des avantages de l'application des modèles à changement de régimes est la prise en compte de la non-normalité des distributions des processus à modéliser. De même, ce modèle permet de tenir compte de l'instabilité temporelle de la relation entre les taux (en particulier l'inflation) et les rendements des actions dont l'origine peut être attribuée à l'émergence de bulles spéculatives et/ou à des fluctuations non dictées par les fondamentaux (cf. Fama [1981]).

A ce niveau, notre étude se distingue par l'adoption d'une démarche novatrice dont l'objectif est de tenir compte de l'instabilité temporelle de la corrélation entre les rentabilités des actifs. Autrement dit, nous tenons compte dans la modélisation des possibilités d'évolution de cette corrélation selon différents régimes.

L'idée principale est de partir de l'approche classique de changement de régime, qui concerne souvent les rendements moyens ainsi que les volatilités, pour supposer que les corrélations basculent elles aussi d'un régime à l'autre. Nous introduisons un nouveau concept qui est celui de la « **corrélation à risque** » correspondant à la matrice de corrélation dans le régime de crise (régime à forte volatilité des marchés). Elle reflètera *a priori* la perception subjective de l'investisseur en fonction des risques auxquels il est exposé. La matrice de corrélation dans le régime à faible volatilité sera considérée comme celle reflétant la stabilité des marchés (estimée par exemple à partir de données historiques hors périodes de crises).

Concernant les valeurs des obligations non gouvernementales (à risque de crédit), ces dernières sont déterminées à travers l'approche des coefficients d'abattement : approche développée par les assureurs dans le cadre de la projection de leurs évaluations de provisions économiques (MCEV[9] et Solvabilité II). Le monétaire est modélisé selon un modèle log-normal à volatilité faible. L'immobilier est modélisé selon un modèle de Vasicek [1977] (l'équivalent de Hull et White [1990] avec un seul facteur : le taux court).

La présentation du GSE, objet de l'illustration, sera accompagnée par la comparaison de ce dernier avec le modèle d'Ahlgrim et al. [2005] notamment au niveau du calibrage des paramètres.

[9] *Market Consistent Embedded Value* est le standard de valorisation des compagnies d'assurance en juste valeur.

I-2 Composantes

En matière de gestion actif-passif, il s'avère utile de s'appuyer sur une modélisation des principales classes d'actifs disponibles dans le but de construire un modèle global dans lequel ces classes sont représentées de manière cohérentes dans une perspective de long terme en phase avec les besoins d'un régime de retraite.

I-2-1 Structure de corrélation

Deux principales questions s'imposent lors de l'étude des modèles de corrélation. D'une part, quelle est l'importance de l'estimation de la corrélation dans le contexte de l'application considérée? D'autre part, quel est le meilleur modèle d'estimation de cette corrélation ?

Concernant la première question sur l'importance de l'estimation des corrélations, nous nous référons à l'étude de Skintzi et al. [2007]. Ces derniers explorent l'importance de l'estimation de la corrélation dans le contexte de la gestion des risques. En particulier, ils étudient l'effet de l'erreur d'estimation de la matrice de corrélation sur le niveau estimé de la Valeur à Risque [10] (VaR ou *Value-at-Risk*) d'un portefeuille d'actifs donné. Ils constatent que l'erreur d'estimation de la matrice de corrélation influence significativement le calcul de la VaR et mettent ainsi en évidence l'importance de minimiser cette erreur pour les gestionnaires de risque (*risk managers*).

Plus précisément, plus l'erreur d'estimation de la matrice de corrélation augmente plus l'erreur au niveau de la VaR augmente. Ceci est vrai que ce soit avec des portefeuilles linéaires (composés d'actifs classiques) ou avec des portefeuilles composés d'options. Pour ces deux portefeuilles, nous supposerons que la VaR est calculée respectivement avec la matrice de variances-covariances et les méthodes de simulation *Monte Carlo*. Ainsi le choix d'un modèle qui minimise l'erreur d'estimation des corrélations est important pour le gestionnaire des risques.

Un résultat additionnel obtenu par les auteurs consiste dans le fait que plus le « vrai » niveau de la corrélation est faible plus la sensibilité du pourcentage d'erreur de calcul de la VaR est élevé. Ceci implique que, dans le cas d'un portefeuille linéaire, la sensibilité est d'autant plus importante que le portefeuille est bien plus diversifié. De même, dans le cas d'un portefeuille d'options, la sensibilité de l'erreur de calcul de la VaR par rapport à l'erreur d'estimation des corrélations est d'autant plus forte que le niveau des corrélations entre les actifs sous-jacents est faible.

Par ailleurs, l'étude de Skintzi et al. [2007] montre que, dans le cas d'un portefeuille d'options, le biais induit par l'erreur d'estimation des corrélations dépend du niveau par rapport à la monnaie et de la maturité. Les résultats montrent à ce niveau que la VaR relative à un portefeuille d'options avec des maturités courtes et *dans la monnaie* est la plus sensible à l'erreur de corrélation. Bien évidemment, ce résultat n'est pas le bienvenu chez les gestionnaires de risque, du fait que la plupart de l'activité de transaction des options est concentré sur des options à maturité courte et proche de la monnaie.

[10] La VaR d'un portefeuille d'actifs financiers correspond au montant de pertes maximales sur un horizon de temps donné et avec un niveau de confiance prédéfini. Il s'agit la mesure de risque de référence utilisée dans Solvabilité II.

Finalement, les résultats de Skintzi et al. [2007] indiquent que la sensibilité de la VaR à l'erreur d'estimation de la matrice de corrélation dépend également de la méthode de calcul de la VaR. Ainsi, la pertinence des résultats obtenus par les simulations *Monte Carlo* doit être vérifiée à travers différents tests de *backtesting*. De même, il est souvent admis que, pour les portefeuilles linéaires, le gestionnaire de risque est indifférent dans le choix entre la matrice de variances-covariances et les simulations *Monte Carlo* : la VaR calculée à travers les méthodes *Monte Carlo* devra converger en principe vers celle obtenue par la matrice de variances-covariances. Cependant, comme le biais provenant de l'erreur d'estimation des corrélations dépend de la méthode employée, l'utilisation des simulations *Monte Carlo* est de préférence à éviter même dans le cas de petits portefeuilles linéaires.

Skintzi et al. [2007] confirment également que quelle que soit l'approche d'estimation retenue, l'erreur est toujours inévitable. En particulier, dans les périodes de stress des marchés, où la VaR est supposé être utile, les corrélations dévient significativement de leurs valeurs estimés initialement (voir Bhansali et Wise [2001] qui mettent en évidence les limites des techniques statistiques dans ce cas). A ce niveau, la particularité de l'étude menée par Skintzi et al. [2007] consiste dans le fait que ces derniers se sont intéressés à la relation entre l'erreur d'estimation de la matrice de corrélation et l'erreur de calcul de la VaR plutôt que de s'intéresser au modèle qui minimise l'erreur d'estimation de la matrice de corrélation.

Autrement dit, ils ont supposé que ce qui compte finalement pour les gestionnaires de risque est de minimiser l'erreur de calcul de la VaR, quel que soit le niveau d'erreur obtenu dans l'estimation de la matrice de corrélation. Ainsi, ils partent de la question : est ce que le choix d'une approche d'estimation de la matrice de corrélation, plutôt qu'une autre approche, impacte le niveau de la VaR obtenu *in fine* et à quel niveau ? Si de telles relations systématiques n'existent pas, le gestionnaire de risque se trouve libre dans le choix de son modèle d'estimation de la matrice de corrélation. Il pourra ensuite, par exemple, chercher à identifier d'autres facteurs menant à des erreurs dans son modèle de calcul de la VaR.

Même si l'importance des hypothèses de corrélation est de premier plan dans certaines applications, du fait de leur influence sur les décisions finales du gestionnaire (cf. Wainscott [1990]), certains chercheurs ont tendance à associer un degré plus faible à cette importance. C'est le cas de Black&Litterman [1992] qui affirment, que dans le cas d'une allocation d'actifs mono-périodique type Markowitz, le modèle est significativement plus sensible aux hypothèses émises sur les espérances et les volatilités des différentes classes d'actifs plutôt qu'aux hypothèses de corrélations adoptées dans le modèle. L'importance des hypothèses de corrélation est donc fonction du contexte d'application et dépend des spécifiés de l'étude envisagée.

Concernant la deuxième question sur la performance de la méthode d'estimation des corrélations, de nombreux travaux ont essayé de trouver la réponse en évaluant la performance de la prévision de différents modèles de corrélations dans le cas de multiples applications financières. Citons par exemple, Elton et Gruber [1973] et Chan et al. [1999] dans le contexte de l'allocation d'actifs, Beder [1995] et Alexander et Leigh [1997] dans le contexte de la gestion des risques, Gibson et Boyer [1998] et Byström [2002] dans le contexte du *pricing* des options. Cependant, l'évidence en matière de la sélection du meilleur modèle de corrélation est loin d'être tranchée. De même, du fait que la plupart des applications nécessite la prévision simultanée de la volatilité et de la corrélation, l'évaluation de la performance de la prévision de la corrélation indépendamment de celle de la volatilité apparaît comme peu cohérente.

La corrélation entre les taux (en particulier l'inflation) et les rendements des actions sera étudiée dans la suite. Ce sujet se caractérise par le fait d'être controversé dans la littérature. L'équation de Fisher [1930] stipule que le taux nominal de rentabilité d'un actif financier, tel que les actions, est égal à la somme de l'inflation anticipée et du taux réel de rentabilité de l'action. Cette identité est basée sur deux hypothèses. La première est relative à l'efficience du marché des actions; tandis que la seconde stipule que le taux réel de rentabilité est déterminé par des facteurs réels, et qu'ainsi il est indépendant des anticipations inflationnistes.

Or, de multiples travaux empiriques (dont Fama [1981] et Kim [2003]) destinés à tester l'équation de Fisher révèlent une relation négative entre la rentabilité des valeurs et l'inflation. Cette nouvelle relation est qualifiée dans la littérature économique de "*stock return-inflation puzzle*". Une multitude d'études fut consacrée à l'analyse de cette relation inverse entre la rentabilité des actifs et une variété de mesures de l'inflation ou de l'inflation anticipée.

L'une des hypothèses implicites à la relation de Fisher est que les marchés financiers sont efficients. L'adoption d'une telle hypothèse est synonyme d'une relation stable entre les prix des actifs financiers et leurs déterminants fondamentaux. Or, les travaux de Shiller [1981] dont les résultats révélaient que la variance des cours boursiers est trop élevée par rapport à celles des fondamentaux vont à l'encontre de l'hypothèse d'efficience des marchés. Artus [1998] affirmait que la relation entre l'inflation et le prix des actifs peut être instable.

L'absence d'une relation stable, au moins à court terme, entre les cours boursiers et leurs déterminants fondamentaux peut être expliquée au niveau macroéconomique, soit par l'irrationalité ou la myopie des agents économiques, soit par l'existence de bulles rationnelles. Autrement dit, étudier la relation entre les prix des actifs financiers et l'inflation sans tenir compte de ces imperfections et anomalies conduit nécessairement à des relations erronées. Le graphique 11 illustre cette instabilité au niveau des corrélations glissantes entre l'inflation et les rendements des actions aux Etats-Unis : par exemple, pour deux crises différentes (celle de 2001 et de celle de 2008) ce niveau de corrélation passe de – 0,2 à 0,2 ce qui s'explique par la différence de la source de la crise et de ses effets sur le comportement des investisseurs.

En adoptant un modèle issu de la combinaison de la théorie de la demande de la monnaie et de la théorie quantitative de la monnaie, Fama [1981] affirmait que la relation négative entre les taux nominaux de rentabilité des actions et l'inflation n'est que le reflet du lien négatif entre cette dernière et l'activité économique réelle. Il explique que dans la mesure où l'activité économique réelle est négativement corrélée à l'inflation et puisque la rentabilité des actions est corrélée positivement à l'activité économique, la corrélation négative entre l'inflation et les taux nominaux de la rentabilité des actions est fausse. Elle ne représente qu'une relation proxy du lien entre les évolutions des prix des actions et de la production.

Ainsi, la thèse dite de "*proxy effect hypothesis*" de Fama est souvent avancée pour justifier les résultats empiriques opposés à l'identité de Fisher. Il faut souligner que la plupart des travaux empiriques relatifs à l'analyse de la relation de l'inflation et la rentabilité des actions ne tiennent pas compte de l'instabilité temporelle de cette relation dont l'origine peut être attribuée à l'émergence de bulles spéculatives et/ou à des fluctuations non dictées par les fondamentaux.

L'instabilité dans le temps de la corrélation entre les différents facteurs de risque, quel que soit l'horizon de calcul de cette corrélation (5 ans, 10 ans, etc.), constitue un élément d'étude important au niveau du choix du niveau de finesse d'un modèle d'actifs (cas de l'allocation

d'actifs). Cette importance provient de l'impact significatif des hypothèses de corrélation retenues sur les résultats finaux obtenus et ainsi sur la décision finale du *manager*.

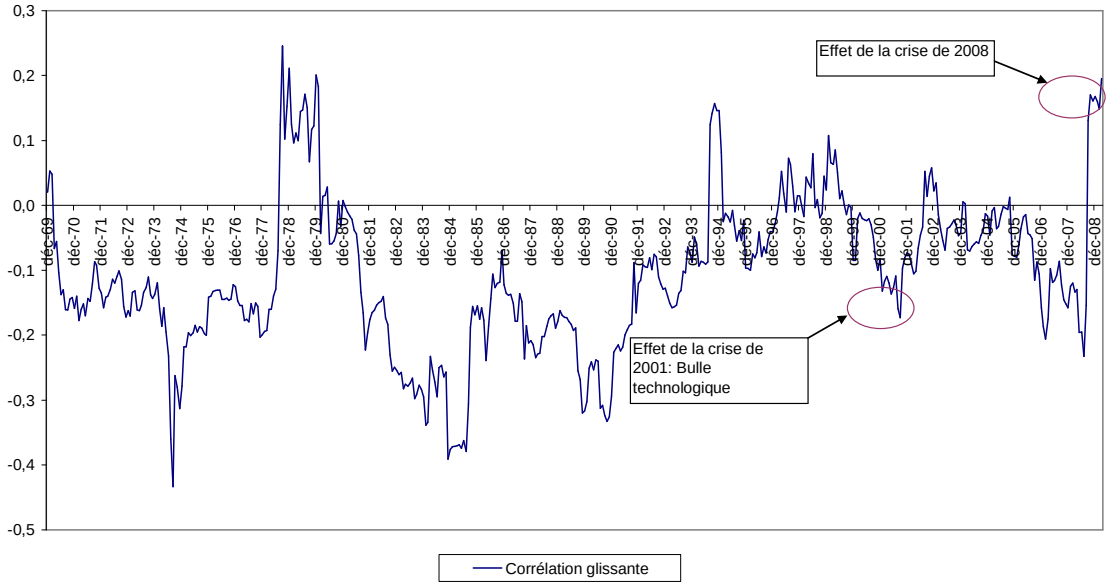


Fig.11 : *Evolution de la corrélation (glissante) sur 5 ans entre l'inflation et le rendement des actions aux Etats-Unis depuis 1970*

L'omission des régimes qui caractérisent les processus générateurs de la rentabilité des actions dans l'analyse de la relation " inflation-rentabilité " peut introduire des biais et suggérer la présence d'une corrélation négative. Si l'introduction des régimes génère une relation conforme à l'identité de Fisher (corrélation positive entre l'inflation anticipée et la rentabilité des actions), nous pourrions mettre en doute les arguments avancés par Fama. Et le rendement des actions serait, alors, un indicateur avancé de l'inflation future.

A ce niveau, notre étude se distingue par l'adoption d'une démarche novatrice dont l'objectif est de tenir compte de l'instabilité temporelle de la corrélation entre les rentabilités des actifs. Autrement dit, nous tenons compte dans la modélisation des possibilités d'évolution de cette corrélation selon différents régimes.

L'idée principale est de partir de l'approche classique de changement de régime, qui concerne souvent les rendements moyens ainsi que les volatilités, pour supposer que les corrélations basculent elles aussi d'un régime à l'autre. Un nouveau concept est ainsi introduit et sera appelé la « **corrélation à risque** ». Il s'agit de la matrice de corrélation dans le régime de crise (régime à forte volatilité des marchés). Elle reflètera *a priori* la perception subjective de l'investisseur en fonction des risques auxquels il est exposé. La matrice de corrélation dans le régime à faible volatilité sera considérée comme celle reflétant la stabilité des marchés (estimée par exemple à partir de données historiques hors périodes de crise).

En fait, pour le régime de faible volatilité (régime 1 pour le rendement des actions), nous avons supposé une matrice de corrélation, entre les variables économiques et financières (le rendement des actions, le rendement de l'immobilier, le monétaire, les taux d'inflation anticipés longs et courts, les taux réels longs et courts) qui est estimée à partir des données historiques après élimination des périodes aberrantes ou de crises. Elle correspond alors à la

matrice de corrélation en situation de stabilité économique. Cependant, nous avons choisi de fixer des corrélations dans le cas de situations de crises (régime 2 pour les rendements des actions), qui correspondent aux corrélations « à risque » pour l'investisseur : il s'agit de la relation entre les variables qui reflètera les craintes de l'investisseur. Ainsi, nous mettons en relief l'exposition au risque extrême en cas de crise.

Si nous considérons par exemple que notre investisseur craint une baisse des actions accompagnée par une hausse des taux (baisse de la valeur des obligations) due à une hausse des taux d'inflation (cas d'un passif indexé à l'inflation) et une baisse des taux réels (cas de ralentissement de l'économie), alors les corrélations sont fixées en situation de crise (régime 2 pour les actions) de façon à avoir une corrélation négative élevée entre les actions et les taux d'inflation, une corrélation positive entre les actions et les taux réels et implicitement une corrélation négative entre les taux d'inflation et les taux réels.

Dans un cadre d'évaluation *Mark-to Market* des engagements, comme la duration du passif est élevée par rapport à celle de l'actif, ce qui est le cas en général, du fait que les maturités des produits obligataires sont relativement courtes, la crainte de l'investisseur sera la baisse des actions accompagnée par une baisse des taux d'inflation anticipés et des taux réels : ce qui ramène le fonds de retraite à une situation de sous financement. Une fois que ce scénario est choisi comme scénario catastrophe, nous fixerons des corrélations positives élevées entre ces variables pour le cas d'un régime à volatilité élevée des actions.

Cette démarche intègre ainsi la méthodologie déterministe de « *stress scenario* » dans un contexte de simulation stochastique. Cela permet, même si la démarche proposée ne présente pas assez d'éléments quantitatifs, de tenir compte dans les simulations de la perception subjective du risque extrême par l'investisseur.

La corrélation entre les actifs est présentée ici dans un cadre **« non linéaire »** (avec deux régimes de corrélation). Au niveau de chaque régime, les corrélations entre les différents actifs du portefeuille peuvent être illustrées par exemple à travers la matrice ci-dessous :

	$dz_t^{OBLIGATION}$	dz_t^{ACTION}	$dz_t^{IMMOBILIER}$	$dz_t^{MONETAIRE}$
$dz_t^{OBLIGATION}$	dt	$\rho_{12}dt$	$\rho_{13}dt$	$\rho_{14}dt$
dz_t^{ACTION}	$\rho_{12}dt$	dt	$\rho_{23}dt$	$\rho_{24}dt$
$dz_t^{IMMOBILIER}$	$\rho_{13}dt$	$\rho_{23}dt$	dt	$\rho_{34}dt$
$dz_t^{MONETAIRE}$	$\rho_{14}dt$	$\rho_{24}dt$	$\rho_{34}dt$	dt

Tab. 6 : *Illustration de la matrice de corrélation entre les différents actifs du portefeuille*

où ρ_{12} : la corrélation entre les obligations et les actions.

Notons, cependant, que les difficultés liées à l'agrégation des risques peuvent être également résolues par l'introduction de copules : ces fonctions permettent de modéliser l'intégralité de la structure de dépendance entre plusieurs variables et donc de prendre en compte les dépendances non linéaires (cf. Horta et al. [2008], Armel et al. [2010]).

I-2-2 Choix des modèles de diffusions

Dans cette étude, le choix d'un modèle de diffusion concerne les variables financières suivantes : la structure par terme des taux d'intérêt (y compris les taux nominaux, réels et d'inflation), le rendement des actions, le rendement des obligations crédits et le rendement de l'immobilier. Ces variables sont retenues compte tenu de leur importance dans le cadre d'une allocation stratégique de long terme d'un fonds de retraite (cf. Campbell et al. [2001]).

I-2-2-1 *Le modèle de structure par terme des taux d'intérêt*

Une attention particulière doit être portée à l'inflation, qui est utilisée pour revaloriser les cotisations et les prestations de tels régimes. Par ailleurs, un modèle d'actifs dans un contexte de retraite se doit d'une part d'intégrer des fluctuations de court terme sur la valeur des actifs et d'autre part d'être cohérent à long terme avec les équilibres macro-économiques reliant l'inflation, les taux d'intérêt et le taux de croissance de l'économie.

En effet, à long terme, il est généralement considéré (cf. Planchet et Thérond [2007]) qu'il existe un lien étroit entre le taux d'intérêt nominal prévalant sur les marchés financiers, le taux d'inflation, c'est-à-dire l'évolution de l'indice des prix, et le taux d'intérêt réel. Plus précisément, le taux d'intérêt nominal est égal (ou converge) à la somme du taux d'intérêt réel et du taux d'inflation :

$$Taux\ d'intérêt\ nominal \approx Taux\ d'inflation + Taux\ d'intérêt\ réel$$

Simultanément, il existe un lien étroit à long terme entre le taux d'intérêt réel et le taux de croissance réel de l'économie. Nous considérons ainsi que le taux d'intérêt réel ne peut être longuement très différent du taux de croissance de l'économie sauf à provoquer des arbitrages entre activité financière et activité réelle d'une part, entre investissement dans un pays et investissement dans d'autres pays d'autre part.

Nous pouvons donc écrire :

$$Taux\ d'intérêt\ réel\ à\ long\ terme \approx Taux\ de\ croissance\ de\ l'économie\ à\ long\ terme$$

Ces contraintes ont donc été prises en compte dans le choix du modèle. Le graphique 12 montre une corrélation forte entre l'inflation et les taux d'intérêt nominaux de court terme au Royaume-Uni à partir de 1960.

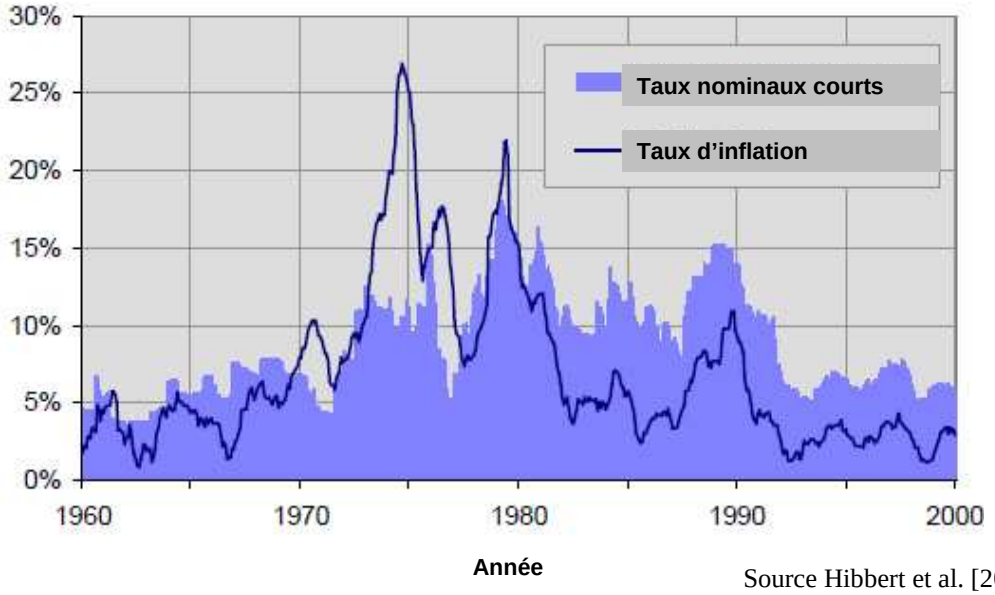
L'idée de base dans le modèle construit : la structure par terme des taux d'intérêt nominaux est établie à partir de deux composantes séparées :

- La structure par terme des taux réels : nous supposons qu'il s'agit de la courbe des taux des obligations indexées à l'inflation.
- Un modèle de l'inflation anticipée pour différents horizons. Il est implicitement supposé que les investisseurs comprennent le processus générant l'inflation et ajustent leurs anticipations de façon cohérente avec la trajectoire déjà déterminée.

Les deux structures par terme sont combinées afin d'obtenir celle des taux nominaux en tenant compte de :

- La corrélation entre ces deux composantes,

- La prime de risque du temps (prime de terme) et/ou la prime de risque inflation.



Source Hibbert et al. [2001]

Fig. 12 : *La corrélation entre les taux d'inflation et les taux nominaux courts au Royaume-Uni*

i. Taux d'intérêt réel : Modèle de Hull-White à 2 facteurs

Nous utilisons un modèle qui est l'extension de l'un des premiers modèles stochastiques de taux d'intérêt : le modèle de Vasicek [1977]. Ce dernier spécifie un processus stochastique en temps continu avec une propriété de retour vers la moyenne pour les taux de court terme (ou les taux instantanés). Nous pouvons ajouter une prime de risque demandée par l'investisseur pour avoir détenu l'obligation plus longtemps que le monétaire. Ce modèle est élargi par l'addition d'un deuxième facteur stochastique. Ce facteur prévoit que le niveau des taux d'intérêt de long terme suit lui aussi un processus stochastique de retour à la moyenne. Ce modèle est inspiré d'un modèle général de structure par terme à 2 facteurs décrit par Hull et White [1994]. Les équations régissant ce processus sont les suivantes :

$$dl_r(t) = k_1(\mu_{lr} - l_r(t))dt + \sigma_1 dZ_1(t)$$
$$dr(t) = k_2(l_r(t) - r(t))dt + \sigma_2 dZ_2(t)$$

Avec :

$r(t)$ = le taux d'intérêt réel à court terme (à l'instant *t*) ou le taux réel instantané.

$l_r(t)$ = le taux d'intérêt réel à long terme (à l'instant *t*).

k_2 = le paramètre d'auto régression du processus des taux réels à court terme (vitesse de retour à la moyenne).

k_1 = le paramètre d'auto régression du processus des taux réels à long terme.

σ_2 = la volatilité annualisée (écart-type) des taux d'intérêt réels à court terme.

σ_1 = la volatilité annualisée (écart-type) des taux d'intérêt réels à long terme.

μ_{lr} = le taux d'intérêt réel à long terme moyen (ou le niveau de retour à la moyenne de $l_r(t)$).

$dZ_2(t)$ = le choc au processus de taux réels à court terme avec une distribution $N(g_r dt, dt)$.

$dZ_1(t)$= le choc au processus de retour à la moyenne des taux réels à court terme avec une distribution $N(g_r dt, dt)$.

g_r = un paramètre qui reflète l'excès de rendement entre les obligations indexées à l'inflation de court et de long terme.

lb_r = la borne inférieure du taux réel à court terme $r(t)$.

hb_r = la borne supérieure du taux réel à court terme $r(t)$.

Le graphique suivant donne une idée sur les résultats pouvant être obtenus par ce modèle. Ici $r(t)$ représente le taux *forward* instantané (court terme). Nous pouvons ensuite déduire la courbe des taux spots en combinant les taux *forwards* appropriés.

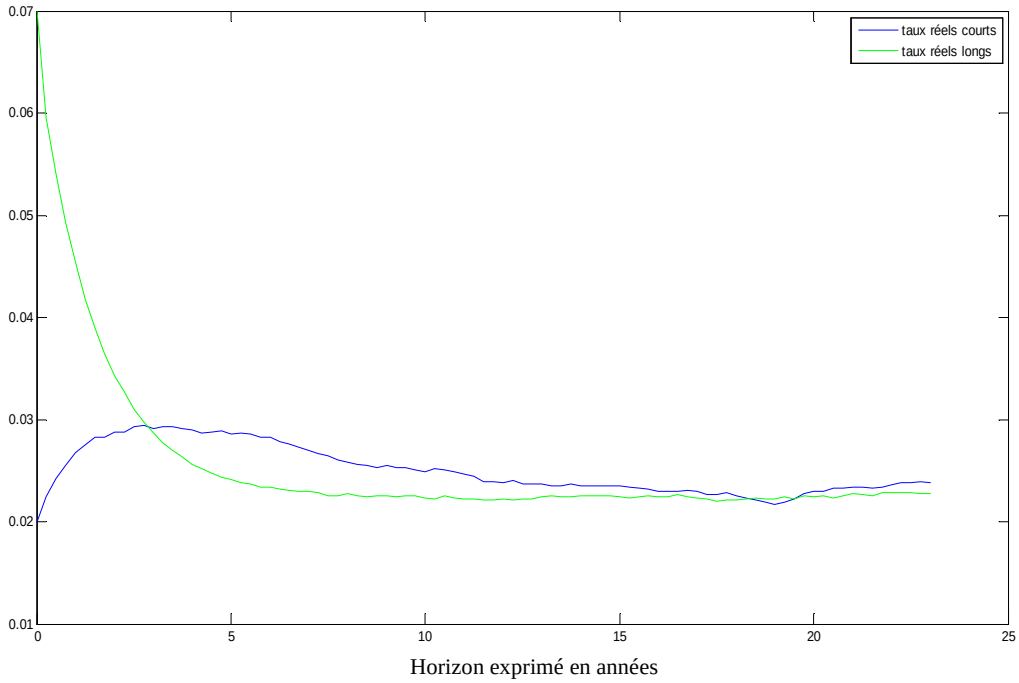


Fig .13 : *Illustration du phénomène de retour à la moyenne des taux réels dans le cadre du modèle de Hull et White à deux facteurs*

Le prix, à l'instant t, d'une obligation zéro-coupon qui paie une unité en terme réel à l'instant T est donné par les équations suivantes (cf. Hibbert et al. [2001]) :

$$P_{réel}(t,T) = exp\left[A(T-t) - B_2(T-t)r(t) - B_1(T-t)l_r(t)\right]$$

Avec :

$$B_2(s) = \left[\frac{1-e^{-k_2 s}}{k_2}\right]$$

$$B_1(s) = \frac{k_2}{k_2 - k_1}\left[\frac{1-e^{-k_1 s}}{k_1} - \frac{1-e^{-k_2 s}}{k_2}\right]$$

$$A(s)=(B_2(s)-s)\left(\mu_{lr}-\frac{\sigma_2^2}{2k_2}\right)+B_1(s)\mu_{lr}-\frac{\sigma_2^2 B_2(s)^2}{4k_2}+\frac{\sigma_1^2}{2}\left[\frac{s}{k_1^2}-2\frac{(B_1(s)+B_2(s))}{k_1^2}+\right.$$

$$\left.\frac{1}{(k_2-k_1)^2}\frac{\left(1-e^{-2k_2 s}\right)}{2k_2}-\frac{2k_2}{k_1(k_2-k_1)}\frac{\left(1-e^{-(k_1+k_2)s}\right)}{(k_1+k_2)}+\frac{k_2^2}{k_1^2(k_2-k_1)^2}\frac{\left(1-e^{-2k_1 s}\right)}{2k_1}\right]$$

Nous pouvons aussi calculer le rendement continu composé à l'instant t pour une maturité T :

$$R_r(t,T)=-log\{P_{réel}(t,T)\}/(T-t)$$

ii. Le modèle d'inflation : Modèle de Hull-White à 2 facteurs

La même structure par terme est utilisée :

$$dl_i(t)=k_3(\mu_{li}-l_i(t))dt+\sigma_3 dZ_3(t)$$
$$di(t)=k_4(l_i(t)-i(t))dt+\sigma_4 dZ_4(t)$$

Avec :

$i(t)$ = le taux d'inflation à court terme (à l'instant t) ou le taux d'inflation instantané.

$l_i(t)$ = le taux d'inflation à long terme (à l'instant t).

k_4 = le paramètre d'auto régression du processus des taux d'inflation à court terme (vitesse de retour à la moyenne).

k_3 = le paramètre d'auto régression du processus des taux d'inflation à long terme.

σ_4 = la volatilité annualisée (écart-type) des taux d'inflation à court terme.

σ_3 = la volatilité annualisée (écart-type) des taux d'inflation à long terme.

μ_{li} = le taux d'inflation à long terme moyen (ou le niveau de retour à la moyenne de $l_i(t)$).

$dZ_4(t)$ = le choc au processus de taux d'inflation à court terme avec une distribution $N(g_i dt, dt)$.

$dZ_3(t)$ = le choc au processus de taux d'inflation à long terme avec une distribution $N(g_i dt, dt)$.

g_i = un paramètre qui reflète la prime de risque inflation entre les obligations nominales et les obligations indexées à l'inflation.

lb_i = la borne inférieure du taux d'inflation à court terme $i(t)$.

hb_i = la borne supérieure du taux d'inflation à court terme $i(t)$.

Nous pouvons de la même façon ensuite déduire la structure par terme des anticipations d'inflation à partir du taux d'inflation à court terme $i(t)$ et de la valeur de $l_i(t)$.

Soit alors : $P(t,T)$ = la valeur actualisée d'une obligation à l'instant t qui paie 1 unité à l'instant T et dont l'actualisation se fait par rapport à l'anticipation d'inflation.

Le rendement équivalent sera :

$$R_i(t,T)=-log\{P_{inf\ lation}(t,T)\}/(T-t)$$

Nous pouvons aussi inclure la prime de risque d'inflation dans la structure par terme des anticipations d'inflation. Comme pour les taux réels, deux bornes (supérieure et inférieure) peuvent être fixées par l'investisseur.

iii. La structure par terme des taux nominaux

Hibbert et al. [2001] explicitent la valeur du taux nominal et le prix d'un zéro coupon nominal en fonction de la relation entre les taux réels et les taux d'inflation anticipés dans le cadre d'un modèle de Hull et White à 2 facteurs.

- Cas de l'indépendance entre les taux réels et les taux d'inflation anticipés :

$$\text{Taux nominal} = i + r$$

où :

i est l'inflation (de court terme ou de long terme suivant le cas) ;

r est le taux d'intérêt réel (de court terme ou de long terme suivant le cas) .

L'origine de cette formule, qui n'est qu'une approximation, provient de la proposition de Fisher [1930] qui stipule que le taux d'intérêt nominal i est déduit de la relation suivante :

$$Taux\ nominal = \left[(1+i)\times(1+r)\right]-1$$

Comme conséquence nous obtenons (Hibbert et al. [2001]) :

$$P_{nom}(t,T) = P_{réel}(t,T)\times P_{inf\ lation}(t,T)$$

- Cas de la dépendance entre les taux réels et les taux d'inflation anticipés :

Nous ajoutons une quantité reflétant la covariance entre les deux variables (Hibbert et al. [2001]) :

$$P_{nom}(t,T) = P_{réel}(t,T)\times P_{inf\ lation}(t,T) + \rho.\sqrt{\left[Var\left(exp\{-R_r(t,T)\}\right).Var\left(exp\{-R_i(t,T)\}\right)\right]}$$

Avec :

ρ = la corrélation entre le choc du taux réel instantané et le taux d'inflation anticipé instantané $dZ_2(t)$ et $dZ_4(t)$.

$$R_r(t,T) = \int_t^T r(s)ds$$

$$R_i(t,T) = \int_t^T i(s)ds$$

Afin de calculer cette covariance, nous partons de l'expression de la variance suivante correspondant aux taux réels:

$$Var\left(exp\{-R_r(t,T)\}\right) = exp\{-2E(R_r(t,T)) + Var(R_r(t,T))\}.(exp\{Var(R_r(t,T))\}-1)$$

Avec :

$$E(R_r(t,T)) = \mu_{lr}.(T-t) + x_1 + x_2$$

$$Var(R_r(t,T)) = y_1.(T-t) - 2y_2y_3 - 2y_4y_5 + (1/k_2)y_6y_7 - 2y_8y_9 + y_{10}y_{11}$$

$$x_1 = (r(t) - \mu_{lr})\frac{(1-exp\{-k_2(T-t)\})}{k_2}$$

$$x_2 = \frac{k_2}{k_1-k_2}\frac{(1-exp\{-k_1(T-t)\})}{k_1} - (l_r(t) - \mu_{lr})\frac{(1-exp\{-k_2(T-t)\})}{k_2}$$

$$y_1 = \left(\frac{\sigma_2}{k_2}\right)^2 + \left(\frac{\sigma_1}{k_1}\right)^2$$

$$y_2 = \left(\frac{\sigma_2}{k_2}\right)^2 - \frac{\sigma_{k_1}^2}{k_1(k_2-k_1)}$$

$$y_3 = \frac{1-exp\{-k_2(T-t)\}}{k_2}$$

$$y_4 = \frac{k_2\sigma_1^2}{k_1^2(k_2-k_1)}$$

$$y_5 = \frac{(1-exp\{-k_1(T-t)\})}{k_1}$$

$$y_6 = \sigma_2^2 + \frac{k_2^2\sigma_1^2}{(k_2-k_1)^2}$$

$$y_7 = \frac{(1-exp\{-2k_2(T-t)\})}{2k_2}$$

$$y_8 = \frac{k_2\sigma_1^2}{k_1(k_2-k_1)^2}$$

$$y_9 = \frac{(1-exp\{-(k_2+k_1)(T-t)\})}{k_2+k_1}$$

$$y_{10} = \frac{k_2^2\sigma_2^2}{k_1^2(k_2-k_1)^2}$$

$$y_{11} = \frac{(1 - exp\{-2k_1(T-t)\})}{2k_1}$$

Une expression analogue est utilisée pour la variance correspondant à la structure des taux d'inflation anticipés. Ainsi, à travers la prise en compte du produit de ces deux variances et de la corrélation, nous obtiendrons le terme de covariance qui permettra le calcul du prix d'un zéro-coupon nominal dans le cadre de la dépendance des taux.

iiii. Prise en compte de la prime de terme dans la STTI

La prime de terme (ou de l'horizon) a pour effet d'ajuster la trajectoire du taux d'intérêt de court terme, de façon à ce que la courbe des taux ne soit pas un estimateur biaisé de l'évolution future effective des taux courts (cf. Hibbert et al. [2001]).

Etant données les valeurs actuelles à l'instant t de $r(t)$ et de $l_r(t)$ (supposées ici être respectivement le taux nominal court et long terme), nous pouvons calculer les valeurs espérées et les variances de ces deux variables à la date $T>t$. Elles sont obtenues par les expressions suivantes :

$$E[l_r(T) \setminus l_r(t)] = \mu_{lr} + e^{-k_1(T-t)}(l_r(t) - \mu_{lr})$$

$$E[r(T) \setminus r(t), l_r(t)] = \mu_{lr} + e^{-k_2(T-t)}(r(t) - \mu_{lr}) + \frac{k_2}{k_2 - k_1}\left(e^{-k_1(T-t)} - e^{-k_2(T-t)}\right)(l_r(t) - \mu_{lr})$$

$$Var[l_r(T) \setminus l_r(t)] = \frac{\sigma_1^2}{2k_1}\left(1 - e^{-2k_1(T-t)}\right)$$

$$Var[r(T) \setminus r(t), l_r(t)] = \frac{\sigma_2^2}{2k_2}\left(1 - e^{-2k_2(T-t)}\right) + \frac{\sigma_2^2}{2k_2}\left(1 - e^{-2k_2(T-t)}\right) + \ldots$$

$$\ldots\left(\frac{k_2\sigma_1}{k_2 - k_1}\right)^2\left[\frac{1}{2k_2}\left(1 - e^{-2k_2(T-t)}\right) + \frac{1}{2k_1}\left(1 - e^{-2k_1(T-t)}\right) - \frac{2}{k_2 - k_1}\left(1 - e^{-(k_2+k_1)(T-t)}\right)\right]$$

Comme $r(t)$ et $l_r(t)$ sont normalement distribuées, nous aurons besoin seulement des moments ci-dessus afin de simuler ces distributions et intégrer à la STTI la prime de terme.

Ainsi :

$$\begin{cases} l_r(T) = E[l_r(T) \setminus l_r(t)] + \sqrt{Var[l_r(T) \setminus l_r(t)]}.Z_1(T) \\ r(T) = E[r(T) \setminus r(t), l_r(t)] + \sqrt{Var[r(T) \setminus r(t), l_r(t)]}.Z_2(T) \end{cases}$$

Avec : $Z_1(T)$ et $Z_2(T)$ sont deux mouvement browniens standards indépendants.

Cependant, si nous avons une prime de terme non nulle, g_r, ces équations sont ajustées de la façon suivante :

$$\begin{cases} l_r(T) = E[l_r(T) \setminus l_r(t)] + \sqrt{Var[l_r(T) \setminus l_r(t)]}.\left(Z_1(T) + g_r.\sqrt{(T-t)}\right) \\ r(T) = E[r(T) \setminus r(t), l_r(t)] + \sqrt{Var[r(T) \setminus r(t), l_r(t)]}.\left(Z_2(T) + g_r.\sqrt{(T-t)}\right) \end{cases}$$

La nouvelle courbe obtenue reflètera l'attitude de l'investisseur vis-à-vis du risque. Par exemple, si l'investisseur exige un rendement supplémentaire sur une obligation de long terme, alors g_r est négative et la valeur espérée finale d'une position dynamique dans le monétaire sera inférieure à la valeur espérée finale d'un investissement initial en obligation de maturité par exemple de 10 ans (avec l'hypothèse d'un rebalancement continu permettant de garder la même maturité dans le portefeuille).

La prime de l'horizon peut être exprimée en termes de rendement espéré sur une obligation zéro-coupon (*Yields*), ou en termes de rentabilité espérée (*Returns*). En se référant à Hibbert et al. [2001], nous avons :

Prime de Terme (Rendement)= $- g_r\left(\frac{\sigma_2}{k_2} + \frac{\sigma_1}{k_1}\right) - \frac{1}{2}\left(\frac{\sigma_2^2}{k_2^2} + \frac{\sigma_1^2}{k_1^2}\right)$

Prime de Terme (Rentabilité)= $- g_r\left(\frac{\sigma_2}{k_2} + \frac{\sigma_1}{k_1}\right)$

Notons qu'une prime de terme positive nécessite une valeur de g_r négative.

Dans le modèle de l'inflation, le paramètre de la prime de risque reflète le rendement supplémentaire exigé sur une obligation nominale par rapport à une obligation indexée à l'inflation, et cette « prime de risque inflation » est intégrée de la même façon dans le modèle de taux d'intérêt réel.

I-2-2-2 *Le modèle des actions*

Face aux limites des modèles classiques (hypothèses de normalité, etc.), nous avons opté pour l'utilisation du modèle de changement de régime (RSLN : *Regime Switching Log Normal*, cf. Hardy [2001]) pour le cas du GSE objet de l'illustration.

En effet, une méthode simple pour tenir compte de la volatilité stochastique est de supposer que la volatilité peut prendre une des *K* valeurs discrètes, en passant d'une valeur à l'autre de façon à définir par le modélisateur. Ceci est la base du modèle lognormal de changement de régime. Ce modèle maintient le caractère de la simplicité fourni par le modèle lognormal classique. De même, il permet de refléter les mouvements extrêmes observés sur les marchés.

Le principe de changement de régime a été introduit par Hamilton [1989] qui a décrit un processus de changement de régime autorégressif. Dans Hamilton et Susmel [1994] différents modèles de changement de régime sont analysés, en testant un nombre différent de régime et des modélisations différentes. L'objectif est de modéliser des séries économétriques hebdomadaires. Les deux chercheurs concluent qu'un modèle à volatilité autorégressive conditionnellement hétéroscédastique de type ARCH est le plus adéquat.

Le modèle de changement de régime a été proposé la première fois comme un modèle traduisant la dynamique des actions par Hardy [2001]. L'idée sous jacente à ce modèle est que les marchés peuvent passer d'un instant à l'autre d'un régime à faible volatilité à un régime à

volatilité relativement élevée. Les périodes de volatilité élevée peuvent être induites par exemple par une politique monétaire de court terme ou par une incertitude économique élevée. Ceci permet de générer des distributions de rendements qui sont plus pertinentes par rapport aux distributions empiriques.

La distribution des rendements des actions est ainsi supposée transiter entre différents états de la nature (régimes). Une matrice de transition contrôle les probabilités de passage entre les régimes. En fait, le processus proposé est markovien : la probabilité de changement de régime dépend seulement du régime courant et non pas de la totalité de la trajectoire historique du processus.

I-2-2-2-1 Le modèle initial de changement de régime

Les rendements des actions ont été décomposés en deux parties : rendement des actions hors dividendes et taux de dividende (Hardy [2001] et Ahlgrim et al. [2005]).

i. Rendement des actions hors dividendes

Les rendements (hors dividendes) des actions, notés s_t, sont égaux au taux d'intérêt nominal (taux sans risque), soit $i_t + r_t$, augmenté d'un excès de rendement attribuable à l'appréciation du capital et noté x_t :

$$s_t = i_t + r_t + x_t$$

Hardy [2001] applique le modèle à changement de régime à l'excès de rendement des actions. En outre, Ahlgrim et al. [2005] reprennent la même approche et estiment les six paramètres de ce modèle à partir de deux jeux de données distincts : un premier jeu (*large stocks*) comprenant des séries observées entre 1871 et 2002 et un deuxième jeu (*small stocks*) comprenant des séries observées entre 1926 et 1999.

ii. Taux de dividende

Ahlgrim et al. [2005] proposent de modéliser le taux de dividende, noté y_t, par :

$$d\left(log\left(y_t\right)\right) = \kappa_y\left(\mu_y - log\left(y_t\right)\right)dt + \sigma_y dB_{y,t}$$

où :

κ_y est la vitesse de retour à la moyenne ;

μ_y est le logarithme du taux de dividende moyen.

Le modèle sur le logarithme du taux de dividende est comparable à celui relatif à l'inflation (avec un seul facteur).

Toutefois, une des principales difficultés du sous-modèle sur le taux de dividende est la disponibilité des données. En effet, les séries de données sur les taux de dividende avec un historique suffisamment long ne sont pas disponibles publiquement. Aussi, pour estimer les paramètres de ce sous-modèle, Ahlgrim et al. [2005] utilisent des données privées. Une autre alternative proposée par Ahlgrim et al. [2005] serait de retenir les taux de dividendes sur un nombre limité d'actions.

De ce fait, nous préférerons travailler avec des rendements d'actions déterminés avec une hypothèse de réinvestissement des dividendes.

I-2-2-2-2 Rendement des actions avec dividendes réinvestis : approche alternative

Au regard des difficultés susceptibles d'être rencontrées dans le cadre des spécifications du modèle d'Ahlgrim et al. [2005] (modélisation à partir de l'excès de rendement sur les actions et des taux de dividendes), une alternative est proposée pour modéliser le rendement des actions.

Cette alternative repose sur une modélisation des rendements totaux des actions (et non uniquement l'excès de rendement) avec dividendes réinvestis. En termes de choix de modèle, l'approche à changement de régime d'Hardy [2001] est maintenue.

Soit S_t le cours des actions (dividende réinvesti) à la date t et $Y_t = log(S_{t+1} / S_t)$ le logarithme du rendement des actions pour la t+1^{ième} période. Dans le modèle à changement de régime, nous considérons :

$Y_t \setminus \rho_t \sim N(\mu_{\rho_t}, \sigma^2_{\rho_t})$ où ρ_t représente le régime appliqué à l'intervalle [t, t+1] (avec ρ_t=1 ou 2) ;

$p_{i,j} = Pr[\rho_{t+1} = j \setminus \rho_t = i]$ représente la probabilité de changement de régime (avec i=1 ou 2, j=1 ou 2) ;

$\Theta = \{\mu_1, \mu_2, \sigma_1, \sigma_2, p_{1,2}, p_{2,1}\}$ représente les paramètres à estimer.

Par ailleurs, nous notons que Hardy a testé dans ce cadre le cas des modèles à deux et à trois régimes (K=2,3) et elle constate qu'en ajoutant un troisième régime la modélisation de S_t n'est pas améliorée de façon significative. Garder deux régimes seulement permet en plus la simplification du modèle et une pertinence meilleure lors de l'estimation des paramètres.

Ainsi, dans le modèle de GSE objet de l'illustration, nous avons également supposé l'existence unique de deux régimes. La projection de la trajectoire des rendements entre ces deux régimes se base sur un algorithme de simulation d'une distribution uniforme cumulée. Cette dernière permet de déterminer le régime de volatilité de la période suivante. Les corrélations des actions (au niveau des browniens) avec les autres variables du modèle diffèrent selon les régimes. En cas de régime à volatilité élevée (crise), nous supposerons une corrélation « à risque » pour l'investisseur comme expliqué ci-dessus.

I-2-2-3 *Le modèle des obligations "crédit"*

Portrait et Poncet [2008] précisent que les titres de créance sont généralement affectés par un risque de crédit, qui englobe un risque de défaut (défaillance de l'emprunteur) et un risque de signature (détérioration du crédit de l'emprunteur). Le *spread* de taux traduit ce risque de crédit : il est égal à la différence entre le taux promis (*a priori* observable pour chaque titre) et le taux sans risque.

Dans le détail, le *spread* est composé de l'espérance de baisse du taux dû à la prise en compte du défaut, d'une prime de risque reflétant l'aversion au risque dans le monde réel et d'une prime de liquidité (cf. Portrait et Poncet [2008]). Deux approches peuvent être envisagées *a priori* pour intégrer le risque de crédit au modèle décrit ici :

- la modélisation des probabilités de défaut historiques et leur lien avec les *spreads* observés pour pouvoir calculer le prix des titres présentant une possibilité de défaut ;

- une démarche simplifiée utilisée couramment dans les MCEV (*Market Consistent Embedded Value*) et les états Solvabilité II (QIS ou *Quantitative Impact Studies*, bilan économique) des assureurs vie et basée sur le calibrage d'abattements sur le nominal du titre obligataire.

Ces deux approches sont décrites ci-après, ainsi qu'une troisième approche d'un modèle « mélange » (qui fera l'objet d'une étude comparative avec la démarche des coefficients d'abattement).

I-2-2-3-1 Modélisation des probabilités de défaut

Deux éléments doivent être décrits ici :

- l'impact du défaut sur la détermination du prix du titre ;

- la dynamique des probabilités historiques de défaut.

i. L'impact du défaut sur le prix du titre :

À l'image de l'illustration de Portrait et Poncet [2008], considérons un titre zéro-coupon de durée T quelconque, qui vaut 1 en date 0 et qui promet un flux $X_{max} = e^{\left(r_T^{max} T\right)}$ en date T, avec r^{max} le taux servi par le titre. L'espérance du flux aléatoire X réellement payé par le titre en T s'écrit alors :

$$E^*(X) = \left(1 - \phi_T^*\right) e^{\left(r_T^{max} T\right)} + \phi_T^* \alpha e^{\left(r_T^{max} T\right)}$$

où :

ϕ_T^* représente la probabilité cumulée, en univers risque neutre, que l'émetteur soit en défaut de paiement en date T ;

α représente la fraction de remboursement de sa dette en cas de défaut de paiement en date T (taux de recouvrement).

Par ailleurs, dans un univers risque-neutre la prime de risque est nulle, et en négligeant la prime de liquidité, nous avons :

$$E^*(X) = e^{(rT)}$$

où r représente le taux sans risque. De ces deux dernières égalités, nous déduisons que le *spread* s est égal à :

$$s = r_T^{max} - r = -\frac{1}{T} log\left(1 - \phi_T^*(1 - \alpha)\right)$$

d'où nous déduisons que la probabilité de défaut cumulée entre 0 et T est :

$$\phi_T^* = \frac{1}{1-\alpha}\left(1 - \frac{1}{e^{(sT)}}\right)$$

Cette approche établie donc un lien formel entre la possibilité de défaut et l'existence d'un *spread* par rapport au taux sans risque. Le prix d'une obligation présentant un risque de défaut $P_D(0,T)$ est alors obtenu à partir du prix sans risque de défaut $P(0,T)$ *via* la formule :

$$P_D(0,T) = (1-\phi_T^*)P(0,T) + \phi_T^* \alpha P(0,T).$$

Les *spreads* sont utilisés pour calculer les probabilités de défaut "risque neutre" intervenant dans la formule ci-dessus.

ii. Les probabilités historiques de défaut

Dans le cadre de l'univers risque-neutre utilisé ci-dessus pour la détermination de l'impact sur le prix de la possibilité de défaut, la prime de risque caractérisant l'aversion au risque est nulle. Au niveau de la probabilité de défaut, cette absence de primes de risque est compensée par une surpondération des évènements défavorables et conduit ainsi à une majoration significative de cette probabilité. En d'autres termes, les probabilités de défaut en univers risque-neutre estimées à partir des *spreads* sont nettement supérieures aux probabilités de défaut historiques (cf. Portrait et Poncet [2008] pour une illustration).

Mais dans la modélisation des scénarios économiques il est nécessaire de disposer non seulement du prix des actifs obligataires à chaque date, mais également de tenir compte des probabilités de défaut historiques observées sur le marché. Ces probabilités de défaut devront être cohérentes avec le rating des contreparties et avec la maturité des titres.

Usuellement ces probabilités sont présentées en fonction de la notation de l'émetteur pour un horizon d'un an. La question de la dynamique temporelle de ces probabilités est délicate et fait l'objet d'une abondante littérature (voir par exemple Bluhm et Overbeck [2007]). Afin de ne pas complexifier trop le modèle, il est retenu ici de considérer les probabilités de défaut comme fixes pour un niveau de notation déterminé.

En pratique, nous utiliserons les probabilités de défaut cumulées communiquées par des agences de notations telles que Moody's ou Standard & Poor's. Ces probabilités historiques permettent de simuler des événements de défaut lors de la projection des flux.

I-2-2-3-2 Démarche des coefficients d'abattement

L'approche ci-dessus présente cependant trois défauts importants en pratique :

- l'estimation des paramètres, notamment dans une perspective de long terme, n'est pas simple et de ce fait les modèles ainsi calibrés s'avèrent peu robustes ;
- la manipulation d'une structure par terme des défauts est contraignante en termes de stockage comme de simulation ;
- la simulation du défaut dans le cadre de la projection de l'actif alourdit l'algorithme et s'avère coûteuse en temps de calcul.

C'est pourquoi il est proposé d'utiliser une méthode alternative développée par les assureurs dans le cadre de leurs évaluations de provisions économiques (MCEV et Solvabilité II). Cette méthode consiste simplement à utiliser pour les obligations avec défaut le modèle d'évaluation des obligations sans risque de défaut en abattant le nominal de manière à recaler le prix théorique issu du modèle avec le prix de marché initial.

Le calcul du rendement du crédit repose sur le coefficient d'abattement qui est le rapport entre le prix du marché de l'obligation crédit et le prix théorique de l'obligation zéro coupon nominale d'une même maturité.

$$coeff_abtt = P_crédit / P_nominal$$

Où $coeff_abtt$ est le coefficient d'abattement recherché. De même, par analogie au calcul du prix de l'obligation nominale, le calcul du prix du marché de l'obligation crédit versant un euros à sa maturité est réalisé à partir de l'actualisation des flux au taux actuariel de marché de l'obligation, soit : $P_crédit = \frac{1}{(Tx_crédit + 1)^M}$

où :

$Tx_crédit$: le taux actuariel observé sur le marché des obligations *corporates* ;
$P_crédit$: le prix de l'obligation à l'instant du calcul ;
M : la maturité restante.

Cette modélisation du défaut est déterministe (le coefficient calculé à la date initiale est supposé être constant sur toute la période de projection), mais présente de nombreux avantages pratiques :

- calibrage simple et cohérence avec les prix de marché ;
- pas de simulation du défaut lors des projections, la valeur des titres reflète déjà ce risque ;
- approche légitimée par son statut de « standard de place ».

Elle présente l'inconvénient de ne pas impacter le rendement des titres par l'introduction du défaut, qui se traduit uniquement par un effet de minoration de la valeur.

I-2-2-3-3 Modèle « mélange» de crédit

Ce modèle s'appuie sur la modélisation des *spreads* avec un modèle à changement de régime, selon la logique suivante :

- le *spread* est projeté avec un modèle RSLN2;
- le prix des zéro-coupons est ensuite calculé à partir du taux d'intérêt (projeté par un modèle de Vasicek) majorés de ce *spread* en remplaçant dans la formule fermée du prix du zéro-coupon sans risque de défaut le taux court par le taux court majoré du *spread* ;
- les prix projetés sont ensuite utilisés pour déterminer le taux de rendement ρ_0 et sur ce rendement est appliqué le taux de défaut et de recouvrement pour obtenir finalement :

$$\rho = (1 + \rho_0) \times (1 - \pi \times (1 - d)) - 1$$

avec π la probabilité de défaut historique et *d* le taux de recouvrement.

Autrement, le « spread » de crédit ou l'écart de taux de rendement exigé entre les obligations gouvernementales et les obligations non gouvernementales est modélisé selon un modèle mixte RSLN-B&S tel que décrit ci-dessus. Le « spread » simulé est intégré dans le modèle de taux d'intérêt afin de déterminer le prix d'un zéro-coupon avec risque de crédit. Ce prix ne tient pas compte de la probabilité de défaut : cette dernière est présente lors du calcul de la performance de l'actif et elle est considérée comme constante. De même, le taux de recouvrement, qui reflète la part récupérée en cas de défaut de la créance initiale, est supposé comme étant une constante.

Outre les difficultés pratiques de calibrage des *spreads* à partir de données historiques, ce modèle s'appuie sur deux approximations fortes et délicates à justifier :

- Déterminer le prix d'un titre présentant un risque de défaut utilisant la formule fermée du modèle de base dans laquelle le taux court r_t est remplacé par le taux court majoré du *spread* implique que $x_t = r_t + s_t$ devrait lui-même suivre un modèle de Vasicek, ce qui n'est pas compatible avec l'hypothèse RSLN2 faite sur s_t ;
- La formule $\rho = (1+\rho_0)\times(1-\pi\times(1-d))-1$ doit être calculée avec des estimations « risque neutre » de π et de *d*, ce qui n'est pas l'usage qui en est fait actuellement.

Au surplus, le défaut est introduit deux fois dans ce modèle : la première *via* le *spread* dans la formule de prix et une seconde fois dans le rendement, alors que la justification de l'existence du *spread* est précisément l'existence du défaut, comme nous l'avons discuté au début de cette partie.

I-2-2-4 *Le modèle de l'immobilier*

Ahlgrim et al. [2005] admettent que les rendements issus de l'immobilier, notés $(re)_t$, sont modélisés par Hulle et White à un facteur (l'équivalent de Vasicek [1977]) :

$$d(re)_t = \kappa_{re}\left(\mu_{re} - (re)_t\right)dt + \sigma_{re}dB_{re,t}$$

où :

κ_{re} est la vitesse de retour à la moyenne ;
μ_{re} est le rendement moyen issu de l'immobilier.

I-2-2-5 *Le monétaire*

Le monétaire (*M*) est modélisé selon un modèle lognormal avec un brownien corrélé avec les autres actifs. Il représente l'actif sans risque de court terme.

$$dM_t = \mu_M \times dt + \sigma_M \times dZ_M$$

Avec :
M est le taux de rendement du monétaire
μ_M est le rendement annuel moyen du monétaire

σ_M est la volatilité du rendement du monétaire

dZ_M est un mouvement brownien corrélé avec les autres variables du modèle

Le graphique 14 illustre la structure adoptée dans le GSE proposé et met en évidence le lien entre les différentes variables.

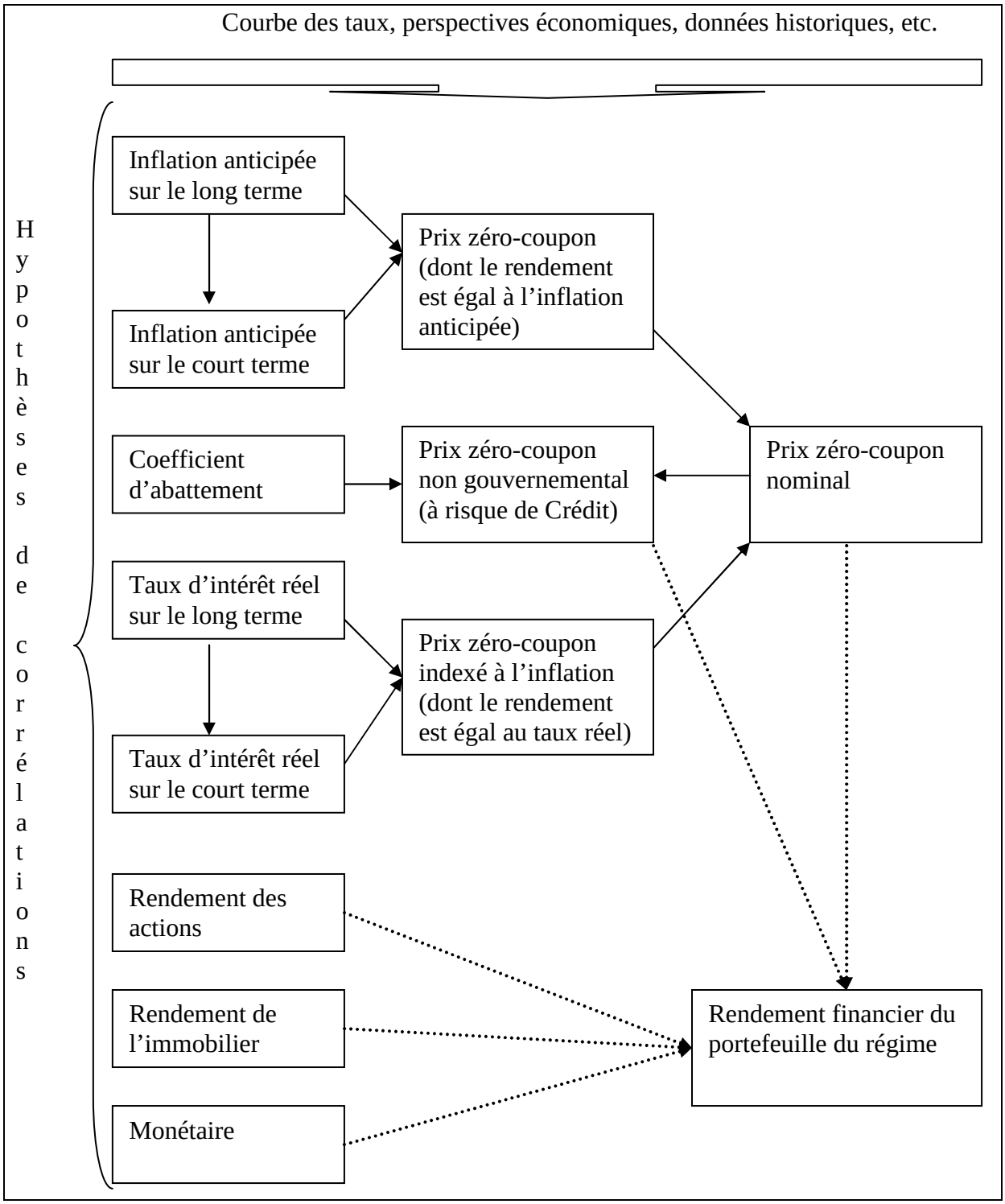


Fig. 14 : *Illustration de la structure du GSE construit*

II- Calibrage

II-1 Présentation du contexte de l'étude

Dans cette sous-section, nous présentons de façon illustrative certaines approches possibles pour le calibrage et l'estimation des paramètres des différents modèles de diffusion. Pour cela il convient de présenter les caractéristiques des données retenues ainsi que les approches générales retenues pour le calibrage.

II-1-1 Description des données retenues

Dans ce contexte, la première étape consiste à identifier les contraintes du modèle. Une attention particulière va être consacrée à l'inflation, principalement en raison du rôle central qu'elle joue dans le modèle de gestion actif-passif.

Par ailleurs, la modélisation de l'actif doit intégrer la perspective de long terme dans laquelle s'inscrit la gestion de cet actif. Cette perspective de long terme a des conséquences au niveau de l'estimation des paramètres du modèle retenu : l'estimation doit être menée sur des séries dont la profondeur doit être cohérente avec l'horizon de projection envisagé. Ainsi, dans le cas de la zone euro par exemple, des données d'un historique de dix ans sont insuffisantes dans le cadre de nos projections de long terme (d'autant plus qu'entre 1999 et 2009 nous comptons deux crises financières dont une particulièrement prononcée en 2008). Plusieurs dizaines d'années d'historique sont ainsi nécessaires pour l'étude.

Outre l'historique, il convient également de s'assurer que la fréquence des données est cohérente avec l'horizon de projection. Les données devront ainsi être de fréquence annuelle, afin de ne conserver que la tendance de long terme des séries.

Les paramètres du GSE proposé seront calibrés, à titre illustratif, sur les données de Friggit [2007] en comparant certains des résultats avec ceux obtenus par Ahlgrim et al. [2005]. Ces données sont présentées ci-après (les données effectivement retenues pour le calibrage sont ***en gras et italique***).

Les données utilisées par Friggit [2007] permettent de répondre à ces contraintes d'historique et de fréquence. Concernant l'historique, l'essentiel des données de Friggit [2007] couvre les années 1800 à 2005 (dans la présente étude, il conviendra toutefois de se limiter aux années 1955 afin de limiter les biais dus à une différence de contexte trop importante – notamment les Guerres mondiales –). Concernant la fréquence, nous noteront que les données de Friggit [2007] sont annuelles.

Pour la zone géographique, on notera que Friggit [2007] présente une série de données pour la France, le Royaume-Uni et les États-Unis. Dans le cadre de la présente étude, seules les données relatives à la France seront retenues pour le calibrage. Aucune donnée relative à la zone euro n'étant disponible dans cette série. Comme mentionné ci-dessus, la disponibilité des données pour la zone euro se limite toutefois à un historique de dix ans.

Séries macroéconomiques	***Indice des prix à la consommation (€, base 2000)***
	Produit intérieur brut (milliards d'€)
	Revenu disponible des ménages (milliards d'€)

Séries démographiques	Population (y compris DOM) (milliers)
	Nombre de ménages, France métropolitaine (milliers)
	Nombre de ménages (y compris DOM) (milliers)

Séries financières	***Taux d'intérêt à long terme***
	Taux d'intérêt à court terme
	Valeur d'un investissement en actions, dividendes réinvestis (€, base 2000)
	Valeur d'un investissement en obligations, intérêts réinvestis (€, base 2000)
	Valeur d'un investissement au marché monétaire, intérêts réinvestis (€, base 2000)
	Prix de l'or (€, base 2000)

Séries immobilières	***Indice du prix des logements, France (€, base 2000)***
	Indice du prix des logements, Paris (€, base 2000)
	Prix "standard" des logements (€) (NB: n'est pas un prix moyen)
	Valeur d'un investissement en logement locatif, Paris (€, base 2000)
	Montants de transactions immobilières (toutes transac-tions soumises aux droits de mutation) (milliards d'€)
	Montants de transactions immobilières (logements anciens) (milliards d'€)
	Nombre de logements commencés (unités)

Tab. 7 : *Les différentes séries retenues pour le calibrage*

Les sources des 5 séries retenues sont les suivantes :

- Indice des prix à la consommation (€, base 2000) : INSEE.
- Taux d'intérêt à long terme : de 1912 à 1959, taux des obligations émises (INSEE), de 1960 à 1988, TMEOG (INSEE) et depuis 19895 : TME (CDC-Ixis).
- Taux d'intérêt à court terme : de 1929 à 1997, taux au jour le jour (INSEE), depuis 1998, TMM (CDC-Ixis).
- Valeur d'un investissement en actions, dividendes réinvestis (€, base 2000) : jusqu'en 1999, (Arbulu [1998]) et depuis 2000: SBF250.
- Indice du prix des logements, France (€, base 2000) : Friggit [2001].

II-1-2 Approche et méthode de calibrage

Il s'agit ici de présenter, de façon générale, l'approche et la méthode de calibrage du modèle. À cet effet, il convient de distinguer principalement deux types de modèle. D'une part les modèles de type Vasicek [1977] qui s'apparentent à des modèles autorégressifs d'ordre 1 après discrétisation (modèles sur l'inflation, les taux d'intérêt réel et les rendements de l'immobilier dans lesquels il convient d'estimer la constante et le coefficient de la variable), et d'autre part les modèles à changement de régime de type Hardy [2001] (modèle sur le rendement total des actions dans lesquels il convient d'estimer pour chaque régime la moyenne, la volatilité et la probabilité de changement de régime).

- Méthode de calibrage des modèles de type autorégressif

Le calibrage des modèles de type autorégressif se fera en plusieurs étapes, décrites ci-après.

› Estimation des paramètres

Dans le cas des modèles de type autorégressif AR(1), l'estimation des paramètres peut être réalisée à partir de la méthode classique des moindres carrés. En particulier, il conviendra de retenir une estimation par les moindres carrés simples pour les modèles de type Vasicek à un-facteur et par les doubles moindres carrés (DMC) pour les modèles de type Vasicek à deux facteurs.

› Tests sur le modèle et les coefficients

La pertinence du modèle sera appréciée à partir de sa qualité d'ajustement et de sa significativité globale. La qualité d'ajustement sera évaluée à partir du coefficient de détermination (R^2 , représente la part de la variance de la variable endogène expliquée par la variance du modèle) et la significativité globale sera mesurée à partir d'un test de Fisher (test dont l'hypothèse nulle H0 correspond à la nullité de tous les coefficients du modèle, cf. Fisher [1954]).

En complément, il conviendra d'évaluer la significativité individuelle des coefficients à partir du test de Student (test dont l'hypothèse nulle H0 correspond à la nullité du coefficient considéré, cf Student [1908]).

Dans les résultats présentés une probabilité est associée aux tests de Fisher et de Student, il s'agit de la probabilité de rejeter à tord H0. Aussi, dans nos analyses, lorsque cette probabilité sera inférieure à 5 %, nous rejetterons H0.

› Analyse des résidus

Une attention particulière doit être accordée à l'analyse des résidus, et dans ce contexte la première étape est de s'assurer que les résidus sont de moyenne nulle.

Dans une seconde étape, il conviendra de détecter l'éventuelle auto-corrélation des erreurs, classique dans les modèles de séries temporelles. À cet effet, au-delà de l'examen visuel des résidus, il conviendra d'utiliser le test de Durbin-Watson pour les auto-corrélations d'ordre 1 (test pour lequel l'hypothèse nulle H0 correspond à l'absence d'auto-corrélation d'ordre 1, cf. Durbin et al. [1951]) et le test de Breusch-Godfrey pour les auto-corrélations d'ordre $p \geq 1$ (test pour lequel l'hypothèse nulle H0 correspond à l'absence d'auto-corrélation d'ordre $p \geq 1$, cf. Godfrey [1979]).

Dans les résultats présentés une probabilité est associée aux tests de Durbin-Watson et de Breusch-Godfrey, il s'agit de la probabilité de rejeter à tord H0. Aussi, dans nos analyses, lorsque cette probabilité sera supérieure à 5 %, nous accepterons H0, c'est-à-dire que nous accepterons l'absence d'auto-corrélation (respectivement d'ordre 1 ou d'ordre $p \geq 1$).

Dans une troisième étape, il conviendra également de détecter une éventuelle hétéroscédasticité, fréquente dans les modèles de séries financières. À cet effet, le test approprié est le test ARCH (*AutoRegressive Conditional Heteroscedasticity*). Ce test permet en effet de modéliser des séries dont la volatilité dépend du passé, et notamment celles présentant de fortes périodes de volatilité suivies de périodes d'accalmies (comme nous en observons dans les séries financières). Ici, l'hypothèse nulle H0 correspond à l'absence de processus ARCH pour justifier le processus.

Dans une quatrième étape, il conviendra de réaliser un test de normalité des erreurs. À cet effet, nous retiendrons le test de Jarque-Bera (cf. Jarque et al. [1980]), fondé sur la notion d'asymétrie (*Skewness*) et d'aplatissement (*Kurtosis*). Pour ce test, l'hypothèse nulle H0 est celle de la normalité des résidus (aussi, lorsque la probabilité associée est supérieure à 5 %, nous acceptons l'hypothèse H0 de normalité des résidus).

- Méthode de calibrage des modèles de type Hardy [2001]

Dans le cas du modèle de type Hardy [2001], les paramètres pourront être estimés à partir de l'approche du maximum de vraisemblance ou de l'approche des filtres de Kalman (détaillées ci-après).

II-2 Calibrage du modèle de l'inflation

Après avoir exposé la méthode retenue pour le calibrage ainsi que les données correspondantes, nous présentons les résultats obtenus des estimations des différents paramètres. Un test d'adéquation est finalement effectué.

II-2-1 Méthode

Pour mémoire, face aux limites constatées sur le modèle d'inflation à un facteur (mauvaise représentation de l'évolution des taux nominaux) l'approche du modèle Hull et White à deux facteurs est retenue.

L'objectif de l'étude envisagée ici ne dépasse pas le cadre de la comparaison avec les résultats obtenus par Ahlgrim et al. [2005] (qui ont adopté le modèle de taux de Hull et White à 1 facteur : les taux courts).

Ainsi, dans le cadre de l'estimation des paramètres du modèle, nous supposerons que la partie court terme i_t est basée sur le modèle à un facteur, avec l'hypothèse simplificatrice que la valeur de long terme de l'inflation intervenant dans la diffusion du processus de court terme est stabilisée à sa valeur limite (qui est égale au taux d'inflation de long terme) :

$$d\,i_t = \kappa_i(\mu_i - i_t)dt + \sigma_i dB_{i,t}$$

soit après discrétisation exacte (cf. Planchet et al. [2005]) :

$$i_{t+1} = i_t \times e^{(-\kappa_i)} + \mu_i\left(1 - e^{(\kappa_i)}\right) + \varepsilon_{i,t} \times \sigma_i \times \sqrt{\frac{1 - e^{(-2\kappa_i)}}{2\kappa_i}}$$

II-2-2 Données

Cette sous-section vise à calibrer un modèle de mesure d'inflation. Cette mesure de l'inflation à la date t est égale à $i_t = log\left(\frac{IPC_t}{IPC_{t-1}}\right)$, où IPC_t représente l'indice des prix à la consommation (disponible en lecture directe dans les données de Friggit [2007]).

II-2-3 Estimation des paramètres

Dans une première étape, l'estimation des paramètres est réalisée à partir d'un modèle autorégressif d'ordre 1 de la forme :

$$i_{t+1} = \alpha_i + \beta_i \times i_t + \varepsilon'_{i,t}$$

Nous obtenons les résultats suivants : $\hat{\alpha}_i = 0{,}01$ et $\hat{\beta}_i = 0{,}77$. Pour revenir au modèle initial décrit par Ahlgrim el al. [2005], les coefficients estimés sont alors transformés. La vitesse de retour à la moyenne est ainsi calculée comme suit :

$$\beta_i = e^{(-\kappa_i)}$$
$$\kappa_i = -log(\beta_i)$$

d'où $\hat{\kappa}_i = 0{,}26$.

De même le taux d'inflation moyen est déterminé par :

$$\alpha_i = \mu_i\left(1 - e^{(-\kappa_i)}\right) = \mu_i\left(1 - e^{(log(\beta_i))}\right)$$
$$\mu_i = \frac{\alpha_i}{(1 - \beta_i)}$$

d'où $\hat{\mu}_i = 0{,}05$.

Enfin, l'écart type de l'erreur du modèle initial est égal à l'écart type du modèle estimé, divisé par $\sqrt{\frac{1-e^{(-2\kappa_i)}}{2\kappa_i}}$. Aussi, l'écart-type du modèle est $\hat{\sigma}_i = 0{,}03$. Le tableau suivant recense les différents paramètres nécessaires à l'alimentation de l'algorithme du modèle d'inflation, en précisant les valeurs obtenues par Ahlgrim et al. [2005] d'une part, et celles obtenues dans le cadre du présent calibrage d'autre part. Nous noterons que le calibrage d'Ahlgrim et al. [2005] est réalisé à partir de données américaines sur la période 1946-2001 (pour mémoire les données de Friggit [2007] retenues dans la présente étude portent sur la France sur la période 1955-2005).

Paramètre	**Intitulé**	**AHLGRIM ET AL.**	**FRIGGIT**
κ_q	Vitesse retour à la moyenne	0,4	**0,26**
μ_q	Taux d'inflation moyen	0,048	**0,05**
σ_q	Écart type de l'erreur	0,04	**0,03**

Tab. 8 : *Calibrage du modèle des taux d'inflation*

Les résultats du calibrage de la présente étude sont comparables à ceux obtenus par Ahlgrim et al. [2005]. Toutefois, il est à noter que les paramètres présentés ne sont pas les paramètres estimés mais sont issus d'une transformation non linéaire de ceux-ci. Cette approche est de nature à introduire un biais dans les paramètres et invite à apprécier la précision des estimateurs avec prudence.

II-2-4 Tests d'adéquation du modèle

Concernant les tests sur le modèle et les coefficients, nous obtenons les résultats suivants :

- coefficient de détermination : $R^2 = 0{,}60$;
- test de significativité globale du modèle (Fisher) : $P_value = 4{,}5.10^{-11}$ (le modèle global est donc significatif) ;
- test de significativité des coefficients (Student) :
    - constante : $P_value = 0{,}04$ (la constante est donc significative)
    - coefficient de la variable : $P_value = 4{,}5.10^{-11}$ (le coefficient de la variable est donc significatif).

Concernant les tests sur l'analyse des résidus, nous obtenons les résultats suivants :

- moyenne des résidus : $2{,}6.10^{-19}$
- test sur l'auto-corrélation d'ordre 1 (Durbin-Watson) : $P_value = 0{,}58$ (absence d'auto-corrélation d'ordre 1)
- test sur l'auto-corrélation d'ordre $p = 1$ (Breusch-Godfrey) : $P_value = 0{,}37$ (absence d'auto-corrélation d'ordre $p = 1$) ;

- test sur l'hétéroscédasticité (ARCH) : $P_value = 0,11$ (absence d'hétéroscédasticité) ;
- test sur la normalité (Jarque-Bera) : $P_value = 2,2.10^{-16}$ (l'hypothèse de normalité des résidus n'est donc pas vérifiée).

Le tableau suivant reprend les principaux résultats sur les tests d'adéquation réalisés.

Test	Résultat	Interprétation
R^2	0,60	Le pouvoir explicatif du modèle est de 60 %
Fisher	$P_value = 4,5.10^{-11}$	Le modèle global est significatif
Student (constante)	$P_value = 0,04$	La constante est significative
Student (variable)	$P_value = 4,5.10^{-11}$	Le coefficient de la variable est donc significatif
Moyenne résidus	$2,6.10^{-19}$	La moyenne des résidus est considérée nulle
Durbin-Watson	$P_value = 0,58$	Absence d'auto-corrélation d'ordre 1
Breusch-Godfrey	$P_value = 0,37$	Absence d'auto-corrélation d'ordre $p = 1$
ARCH	$P_value = 0,11$	Absence d'hétéroscédasticité
Jarque-Bera	$P_value = 2,2.10^{-16}$	Hypothèse de normalité des résidus non vérifiée

Tab. 9 : *Résultat des tests d'adéquation pour le modèle des taux d'inflation*

Hormis le test sur la normalité des résidus, l'ensemble de ces tests valide les conditions posées *a priori* pour la validation du modèle et du calibrage. Nous noterons toutefois que l'hypothèse de normalité des résidus n'étant pas vérifiée, il conviendra de considérer avec prudence d'une part les tests de Student, et d'autre part les intervalles de confiance des projections.

II-3 Calibrage du modèle des taux réels

Après avoir exposé la méthode retenue pour le calibrage ainsi que les données correspondantes, nous présentons les résultats obtenus des estimations des différents paramètres. Un test d'adéquation est finalement effectué.

II-3-1 Méthode

Pour mémoire, le modèle retenu pour les taux d'intérêt réel est le suivant (avec r_t pour les taux courts, et l_t pour les taux longs) :

$$\begin{aligned} dr_t &= \kappa_r \left(l_t - r_t\right) dt + \sigma_r dB_{r,t} \\ dl_t &= \kappa_l \left(\mu_l - l_t\right) dt + \sigma_l dB_{l,t} \end{aligned},$$

soit après discrétisation exacte (cf. Planchet et al. [2005]) :

$$r_{t+1} = r_t \times e^{(-\kappa_r)} + l_t \left(1 - e^{(-\kappa_r)}\right) + \varepsilon_{r,t} \times \sigma_r \times \sqrt{\frac{1 - e^{(-2\kappa_r)}}{2\kappa_r}}$$

$$l_{t+1} = l_t \times e^{(-\kappa_l)} + \mu_l \left(1 - e^{(-\kappa_l)}\right) + \varepsilon_{l,t} \times \sigma_l \times \sqrt{\frac{1 - e^{(-2\kappa_l)}}{2\kappa_l}}$$

Au regard de la dépendance entre les taux long et les taux courts, l'utilisation des Moindres Carrées Ordinaires (MCO) pour l'estimation des modèles n'est pas appropriée. L'estimation des paramètres peut être réalisée à partir d'une application des MCO en deux étapes :

$$\text{Etape 1 : } l_{t+1} = \beta_1 + \beta_2 \times l_t + \varepsilon'_{l,t}$$

$$\text{Etape 2 : } r_{t+1} = \alpha_1 \times \hat{l}_t + \alpha_2 \times r_t + \varepsilon'_{r,t}$$

Pour revenir au modèle initial, les coefficients estimés sont alors transformés : nous déduisons la vitesse de retour à la moyenne des taux réels à long terme ($\kappa_l = -log(\beta_2)$), le taux d'intérêt réel moyen à long terme ($\mu_l = \frac{\beta_1}{(1-\beta_2)}$) et l'écart type de l'erreur du modèle initial de taux à long terme (égal à l'écart type du modèle estimé divisé par $\sqrt{\frac{1-e^{(-2\kappa_l)}}{2\kappa_l}}$)

Une fois les paramètres du premier modèle estimé (étape 1 des DMC), il convient de modéliser les taux d'intérêt réels à court terme, en retenant les taux réels à long terme estimés comme variable explicative (étape 2 des DMC). À la différence des modèles précédant, ce modèle sur les taux courts compte deux variables endogènes :

$$r_{t+1} = \alpha_1 \times \hat{l}_t + \alpha_2 \times r_t + \varepsilon'_{r,t}$$

où les coefficients sont $\alpha_2 = e^{(-\kappa_r)}$ et $\alpha_1 = \left(1 - e^{(-\kappa_r)}\right)$, soit $\alpha_2 = 1 - \alpha_1$. Dans ce contexte, le modèle retenu pour l'estimation des paramètres relatifs aux taux courts est :

$$\Delta r_{t+1} = \alpha_1 \left(\hat{l}_t - r_t\right) + \varepsilon'_{r,t}.$$

Des résultats de ce modèle, nous déduisons la vitesse de « retour à la moyenne » (il s'agit en pratique d'un « retour » vers les taux réels à long terme) des taux réels à court terme ($\kappa_r = -log(1-\alpha_1)$) et l'écart type de l'erreur du modèle initial de taux à court terme (égal à l'écart type du modèle estimé divisé par $\sqrt{\frac{1-e^{(-2\kappa_r)}}{2\kappa_r}}$).

II-3-2 Données

Cette sous-section vise à calibrer les modèles de taux d'intérêt réels (taux long et taux court). Ces taux sont disponibles en lecture directe dans les données de Friggit [2007].

II-3-3 Estimation des paramètres

Estimation des paramètres : 1ère étape des MCO

Concernant la première étape, nous obtenons les résultats suivants : $\hat{\beta}_1 = 0{,}01$ et $\hat{\beta}_2 = 0{,}64$. A partir des équations ci-dessus, il en ressort les résultats suivants : $\hat{k}_l = 0{,}45$, $\hat{\mu}_l = 0{,}03$ et $\hat{\sigma}_l = 0{,}02$.

Estimation des paramètres : 2ème étape des MCO

Nous obtenons $\hat{\alpha}_1 = 0{,}33$. A partir des équations ci-dessus, il en ressort les résultats suivants : $\hat{k}_r = 0{,}40$ et $\hat{\sigma}_r = 0{,}02$.

Le tableau suivant recense les différents paramètres nécessaires à l'alimentation de l'algorithme du modèle de taux d'intérêt réel, en précisant les valeurs obtenues par Ahlgrim et al. [2005] d'une part, et celles obtenues dans le cadre du présent calibrage d'autre part. Nous noterons que le calibrage d'Ahlgrim et al. [2005] est réalisé à partir de données américaines sur la période 1982-2001 (pour mémoire les données de Friggit [2007] retenues dans la présente étude portent sur la France sur la période 1955-2005).

Paramètre	Intitulé	AHLGRIM ET AL.	FRIGGIT
κ_r	Vitesse retour à la moyenne (court terme)	6,1	**0,40**
σ_r	Écart type de l'erreur (court terme)	0,1	**0,02**
κ_l	Vitesse retour à la moyenne (long terme)	5,1	**0,45**
μ_l	Taux d'intérêt réel moyen (long terme)	0,028	**0,03**
σ_l	Écart type de l'erreur (long terme)	0,1	**0,02**

Tab. 10 : *Calibrage du modèle des taux d'intérêt réels*

Malgré un taux d'intérêt réel moyen (long terme) comparable, les résultats du calibrage obtenus lors de l'étude avec les données de Friggit [2007] présente des différences significatives avec ceux obtenus par Ahlgrim et al. [2005].

Ces différences sont toutefois justifiées par la nature des données retenues par Ahlgrim et al. [2005] pour calibrer le modèle : ils retiennent des données mensuelles. Comme indiqué ci-dessus, en retenant des données à fréquence mensuelle nous introduisons un bruit de court terme qui n'est pas cohérent avec les besoins de long terme. Ce bruit se traduit dans les résultats d'Ahlgrim et al. [2005] par une vitesse de retour à la moyenne et par une volatilité de l'erreur supérieures à celles obtenues dans le calibrage résultant des données mensuelles de Friggit [2007].

Nous noterons enfin que là encore les paramètres présentés ne sont pas les paramètres estimés mais sont issus d'une transformation non linéaire de ceux-ci. Cette approche est de nature à introduire un biais dans les paramètres et invite à apprécier la précision des estimateurs avec prudence.

II -3-4 Tests d'adéquation du modèle

Tests d'adéquation du modèle : 1ère étape des MCO

Le tableau suivant reprend les principaux résultats sur les tests d'adéquation réalisés pour la première étape des MCO (modèle sur les taux d'intérêt réels longs).

Test	Résultat	Interprétation
R^2	0,41	Le pouvoir explicatif du modèle est de 41 %
Fisher	$P_value = 6,0.10^{-7}$	Le modèle global est significatif
Student (constante)	$P_value = 0,02$	La constante est significative
Student (variable)	$P_value = 6,0.10^{-7}$	Le coefficient de la variable est donc significatif
Moyenne résidus	$2,8.10^{-19}$	La moyenne des résidus est considérée nulle
Durbin-Watson	$P_value = 0,67$	Absence d'auto-corrélation d'ordre 1
Breusch-Godfrey	$P_value = 0,30$	Absence d'auto-corrélation d'ordre $p = 1$
ARCH	$P_value = 0,72$	Absence d'hétéroscédasticité
Jarque-Bera	$P_value = 2,2.10^{-16}$	Hypothèse de normalité des résidus non vérifiée

Tab. 11 : *Résultats des tests d'adéquation pour le modèle de taux d'intérêt réels (1ère étape)*

Hormis le test sur la normalité des résidus, l'ensemble de ces tests valide les conditions posées a priori pour la validation du modèle et du calibrage. Nous noterons toutefois que l'hypothèse de normalité des résidus n'étant pas vérifiée, il conviendra de considérer avec prudence d'une part les tests de Student, et d'autre part les intervalles de confiance des projections.

Tests d'adéquation du modèle : 2ème étape des MCO

Le tableau suivant reprend les principaux résultats sur les tests d'adéquation réalisés pour la deuxième étape des MCO (modèle sur les taux d'intérêt réels courts).

Test	Résultat	Interprétation
R^2	0,15	Le pouvoir explicatif du modèle est de 15 %
Fisher	$P_value = 6,1.10^{-3}$	Le modèle global est significatif
Student (variable)	$P_value = 6,1.10^{-3}$	Le coefficient de la variable est donc significatif
Moyenne résidus	$3,7.10^{-3}$	La moyenne des résidus est considérée nulle
Durbin-Watson	$P_value = 0,51$	Absence d'auto-corrélation d'ordre 1
Breusch-Godfrey	$P_value = 0,40$	Absence d'auto-corrélation d'ordre $p = 1$
ARCH	$P_value = 0,98$	Absence d'hétéroscédasticité
Jarque-Bera	$P_value = 2,2.10^{-16}$	Hypothèse de normalité des résidus non vérifiée

Tab. 12 : *Résultats des tests d'adéquation pour le modèle de taux d'intérêt réels (2ème étape)*

Hormis le test sur la normalité des résidus, l'ensemble de ces tests valide les conditions posées a priori pour la validation du modèle et du calibrage. Comme pour la première étape, nous noterons que l'hypothèse de normalité des résidus n'étant pas vérifiée, il conviendra de considérer avec prudence d'une part les tests de Student, et d'autre part les intervalles de confiance des projections.

II-4 Calibrage du modèle des actions

Au regard des difficultés pratiques relatives aux spécifications du modèle d'Ahlgrim et al. [2005] pour la modélisation des actions (modélisation de l'excès de rendement sur les actions d'une part et des taux de dividendes d'autre part), une alternative a été proposée pour modéliser le rendement des actions.

Cette alternative repose sur une modélisation des rendements totaux des actions (et non uniquement sur l'excès de rendement) avec dividendes réinvestis par une approche à changement de régime de type Hardy [2001].

II-4-1 Méthode

Méthode 1 : Méthode du Maximum de Vraisemblance

Soit $Y_t = log(S_{t+1} / S_t)$ le log rendement du mois t+1. La vraisemblance des observations $y = (y_1, y_2, ..., y_n)$ est $L(\Theta) = f(y_1 \setminus \Theta) f(y_2 \setminus \Theta, y_1) f(y_3 \setminus \Theta, y_1, y_2) ... f(y_n \setminus \Theta, y_1, ..., y_{n-1})$ Avec f la fonction de distribution de y.

Ainsi la contribution, à la log-vraisemblance, de l'observation de la date t est :

$$log\ f(y_t \setminus y_{t-1}, y_{t-2}, ..., y_1, \Theta)$$

Nous pouvons calculer cette expression de façon récursive (en suivant Hamilton et Sumsel [1994], par exemple), en calculant pour une observation à la date t :

$$f(\rho_t, \rho_{t-1}, y_t \setminus y_{t-1}, ..., y_1, \Theta) = p(\rho_{t-1} \setminus y_{t-1}, ..., y_1, \Theta) \times p(\rho_t \setminus \rho_{t-1}, \Theta) \times f(y_t \setminus \rho_t, \Theta) \qquad (E)$$

où :

$p(\rho_t \setminus \rho_{t-1}, \Theta)$ est la probabilité de transition entre les régimes.

$f(y_t \setminus \rho_t, \Theta) = \phi((y_t - \mu_{\rho_t}) / \sigma_{\rho_t})$ avec ϕ la fonction de densité d'une loi normale standard.

$p(\rho_{t-1} \setminus y_{t-1}, ..., y_1, \Theta)$ est obtenue à partir de l'itération précédente. Elle est égale à :

$$p(\rho_{t-1} \setminus y_{t-1}, ..., y_1, \Theta) = \frac{f(\rho_{t-1}, \rho_{t-2} = 1, y_{t-1} \setminus y_{t-2}, y_{t-3}, ..., y_1, \Theta) + f(\rho_{t-1}, \rho_{t-2} = 2, y_{t-1} \setminus y_{t-2}, y_{t-3}, ..., y_1, \Theta)}{f(y_{t-1} \setminus y_{t-2}, y_{t-3}, ..., y_1, \Theta)}$$

Dans le cas d'un modèle RSLN à deux régimes (ρ=1 ou 2), l'équation (*E*) peut avoir quatre valeurs en fonctions de celles prises par ρ_t et ρ_{t-1}, à savoir :

$$f(\rho_t = 1, \rho_{t-1} = 1, y_t \setminus y_{t-1}, ..., y_1, \Theta)$$
$$f(\rho_t = 1, \rho_{t-1} = 2, y_t \setminus y_{t-1}, ..., y_1, \Theta)$$
$$f(\rho_t = 2, \rho_{t-1} = 1, y_t \setminus y_{t-1}, ..., y_1, \Theta)$$
$$f(\rho_t = 2, \rho_{t-1} = 2, y_t \setminus y_{t-1}, ..., y_1, \Theta)$$

$f(y_{t-1} \setminus y_{t-2}, y_{t-3}, ..., y_1, \Theta)$ **sera ainsi la somme de ces quatre possibilités d'état de la nature.**

Pour commencer l'itération, nous avons besoin d'une valeur initiale pour $p(\rho_0)$. Cette dernière est déduite à partir de la distribution invariante de la chaîne de Markov. En effet, la

distribution invariante $\pi = (\pi_1, \pi_2)$ est la distribution de la probabilité inconditionnelle du processus. Sous cette distribution π, chaque transition est affectée par la même distribution, c'est-à-dire $\pi P = \pi$, étant donné $\pi_1 p_{1,1} + \pi_2 p_{2,1} = \pi_1$ et $\pi_1 p_{1,2} + \pi_2 p_{2,2} = \pi_2$. Autrement $p_{1,1} + p_{1,2} = 1$ et donc $\pi_1 = \frac{p_{2,1}}{p_{1,2} + p_{2,1}}$. De même $\pi_2 = 1 - \pi_1 = \frac{p_{1,2}}{p_{1,2} + p_{2,1}}$.
Ainsi, pour un ensemble de paramètre de départ Θ, nous pouvons commencer le calcul récursif comme suit :

$$f(\rho_1 = 1, y_1 \setminus \Theta) = \pi_1 \phi\left(\frac{y_1 - \mu_1}{\sigma_1}\right)$$
$$f(\rho_1 = 2, y_1 \setminus \Theta) = \pi_2 \phi\left(\frac{y_1 - \mu_2}{\sigma_2}\right)$$
$$f(y_1 \setminus \Theta) = f(\rho_1 = 1, y_1 \setminus \Theta) + f(\rho_1 = 2, y_1 \setminus \Theta)$$

Et nous calculons deux valeurs relatives à l'expression suivante :

$$p(\rho_1 \setminus y_1, \Theta) = \frac{f(\rho_1, y_1 \setminus \Theta)}{f(y_1 \setminus \Theta)} \qquad \rho_1 = 1,2$$

Cette expression sera utilisée dans l'itération suivante lors du calcul de $f(\rho_2, \rho_1, y_2 \setminus y_1, \Theta)$

L'objectif sera donc la maximisation du log de la fonction de vraisemblance résultante.

Autre formulation de l'approche :

Notations :

p_{ij} probabilité de transition de l'état i vers l'état j.

$$p_{ij} = p_{ij}(t) = P(\rho_t = j \setminus \rho_{t-1} = i, \Theta)$$

$\Theta_n(t)$ probabilité d'obtenir l'observation y_t sachant que nous avons dans le régime n.

$$\Theta_n(t) = f(y_t \setminus \rho_t = n, \Theta)$$
$$\Theta_n(t) = \Phi\left(\frac{y_t - \mu_n}{\sigma_n}\right)$$

$Y_{ij}(t)$ probabilité d'être passé du régime i à j et d'observer y_t connaissant l'historique de la série y en $t-1$.

$$Y_{ij}(t) = f(\rho_{t-1} = i, \rho_t = j, y_t \setminus y_{t-1}, ..., y_1, \Theta)$$
$$Y_{\cdot j}(t) = Y_{1j}(t) + Y_{2j}(t)$$
$$Y(t) = Y_{\cdot\cdot}(t) = Y_{\cdot 1}(t) + Y_{\cdot 2}(t)$$

$P_k(t-1)$ Probabilité d'être dans le régime k en $t-1$, sachant l'historique de la série y en $t-1$.

$$P_k(t-1) = P(\rho_{t-1} = k \setminus y_{t-1},...,y_1,\Theta)$$
$$P_k(t-1) = \frac{Y_{1k}(t-1)+Y_{2k}(t-1)}{Y(t-1)} = \frac{Y_{.k}(t-1)}{Y(t-1)}$$

La relation principale est la suivante :

$$Y_{ij}(t) = P_i(t-1)p_{ij}\Phi_j(t)$$
$$Y_{ij}(t) = \frac{Y_{1i}(t-1)+Y_{2i}(t-1)}{Y(t-1)}p_{ij}\Phi_j(t)$$

Pour l'initialisation de l'algorithme :

$$\pi_1 = \frac{p_{21}}{p_{12}+p_{21}}$$
$$\pi_2 = 1-\pi_1$$
$$Y_{.i}(1) = \pi_i\Phi_i(1)$$
$$P_i(1) = \frac{Y_{.i}(1)}{Y(1)}$$

Méthode 2 : Filtre de Kalman

La procédure générale est présentée en détail par Hamilton et al. [1994]. Elle est basée sur la méthode du filtre de Kalman.

- ***Présentation générale de l'approche du filtre de Kalman***

Nous supposons que la série temporelle y_t, représentée par le vecteur $(y_0, y_1, ..., y_n)$ est observable. Cette variable peut par exemple être le taux de rendement d'un titre financier. Elle dépend de la variable x_t qui elle n'est pas observable. Comme l'on ne peut observer x_t, il faudra simuler sa valeur. Nous ne connaissons pas non plus la variance de x_t, représentée par ω_t. Le modèle se présente comme suit :

$y_t = \theta_1 + \theta_2 x_t + \varepsilon_t$ (Equation de mesure)

$x_{t+1} = \theta_3 + \theta_4 x_t + \eta_t$ (Equation de transition ou d'état)

Avec θ_i les paramètres à estimer ; ε_i un bruit gaussien dont la variance est de ν_{1t} et η_t, un bruit gaussien dont la variance est de ν_{2t}. La première équation est dite équation de mesure alors que la seconde est l'équation de transition ou d'état.

Au temps (*t*-1), des estimés de x_{t-1}, de sa variance ω_{t-1} ainsi que des coefficients $\theta_{i,t-1}$ doivent être fournis. Si nous nous situons au temps 0, nous devons disposer d'une estimation préliminaire de x_0 et de ω_0. Mais comme ces valeurs sont alors inconnues, nous pourrons attribuer une valeur nulle à x_0 et une valeur élevée à ω_0 de façon à prendre en compte l'incertitude importante qui est alors reliée à l'estimation de x_0.

Revenons au temps (*t*-1) de la simulation (ou bien du filtrage) et donnons les trois étapes de la procédure suivie par le filtre de Kalman.

L'étape 1: étape de la prévision

Nous calculons alors les deux prévisions suivantes: $x_t / t-1$, la prévision de x_t au temps (*t*-1), soit l'espérance conditionnelle de x_t, étant donnée l'information disponible au temps (*t*-1) ; $\omega_t / t-1$, la prévision de ω_t au temps (*t*-1), soit l'espérance conditionnelle de ω_t, étant donnée l'information disponible au temps (*t*-1). Ces prévisions, qui sont des estimations conditionnelles non biaisées, se calculent comme suit :

L'étape 2: étape de la révision

Au temps *t*, nous disposons d'une nouvelle observation de *y*, soit y_t. Nous pouvons alors calculer l'erreur de prévision υ_t :

$$\upsilon_t = y_t - \theta_{1,t-1} - \theta_{2,t-1} x_{t/t-1}$$

La variance de υ_t , représentée ψ_t par, est de :

$$\psi_t = \theta_{2,t-1}^2 \omega_{t/t-1} + \nu_{1,t-1}$$

Nous nous servons de υ_t et de ψ_t pour mettre à jour x_t et sa variance ω_t :

$$x_t = x_{t/t-1} + \frac{\theta_{2,t-1} \times \omega_{t/t-1} \times \upsilon_t}{\psi_t}$$

$$\omega_1 = \omega_{t/t-1} + \frac{\theta_{2,t-1}{}^2 \times \omega_{t/t-1}{}^2}{\psi_t}$$

Ces deux derniers estimateurs sont les estimateurs non biaisés conditionnellement qui minimisent la variance. Le filtre de Kalman est donc optimal, en ce sens qu'il est le meilleur estimateur dans la classe des estimateurs linéaires (cf. Hamilton et al. [1994]).

L'étape3 : estimation des paramètres

Nous recourons à la méthode de maximum de vraisemblance pour estimer les paramètres θ_i.
La fonction de vraisemblance est la suivante :

$$\ell = -\frac{1}{2}\sum_t log(\psi_t) - \frac{1}{2}\sum_t \frac{\upsilon_t^2}{\psi_t}$$

Nous passons au temps (*t*+1) et nous refaisons cette procédure en trois étapes jusqu'à la période *n*.

- ***Mise en place l'approche du filtre de Kalman pour le RSLN (Chourdakis [2004])***

Soit :

- x_t l'état de la nature (le régime) à l'instant t ($x_t = 1,2,...,N$) qui est une variable non observable et déterminant le rendement financier y_t .

- p_{ij} la probabilité de passage de l'état i à l'état j

- D représente un vecteur des probabilités de passage vers un état défini j. Exemple pour le cas d'une modèle à 2 régimes : $D((1)) = (\ p_{11} \quad p_{21}\ )$

A l'instant *t*, nous partons des probabilités $P_{t-1\backslash t-1}(i) = Pr\{x_{t-1} = i \setminus \Im_{t-1}\}$ et la donnée y_t. $\Im_{t-1}$ est la filtration reflétant les informations sur la trajectoire des régimes jusqu'à $t-1$.

Chourdakis [2004] part de l'approche des filtres de Kalman pour présenter sa propre méthode. Les étapes de cette méthode sont illustrées comme suit :

- La **prévision** de la probabilité de l'état futur : $P_{t\backslash t-1}(i) = D(i).P_{t-1\backslash t-1}(i)$

  Exemple : $P_{t\backslash t-1}(1) = p_{11}.P_{t-1\backslash t-1}(1) + p_{21}.P_{t-1\backslash t-1}(2)$ si nous supposons deux régimes seulement dans notre modèle.

- Le calcul de la densité pondérée de y_t : $\zeta_t(i) = f(y_t \setminus x_t = i)P_{t\backslash t-1}(i)$

- La **filtration** de la probabilité de l'état futur: $P_{t\backslash t}(i) = \dfrac{\zeta_t(i)}{\sum_{j=1}^{N}\zeta_t(j)}$

Nous obtenons ainsi la fonction du log de vraisemblance : $\sum_t log \sum_j^N \zeta_t(j).$

Pour la mise en oeuvre, nous avons opté pour l'approche composée des deux étapes suivantes inspirée de celle de Chourdakis [2004] :

<u>Calcul de la vraisemblance à maximiser</u>

Inputs : Y vecteur des rendements historiques observés (2 lignes x T colonnes)

$$Y_0 = \begin{bmatrix} \dfrac{p_{21}}{p_{21}+p_{12}} \\ \dfrac{p_{12}}{p_{21}+p_{12}} \end{bmatrix}$$

$$dens = \begin{bmatrix} f_Y(Y, m1 - 0{,}5 \times s_1^{\,2}, s_1) \\ f_Y(Y, m2 - 0{,}5 \times s_2^{\,2}, s_2) \end{bmatrix} \qquad \text{avec } f_Y \text{ : densité de la loi normale}$$

$$cp = \begin{bmatrix} 1-p_{12} & p_{21} \\ p_{12} & 1-p_{21} \end{bmatrix}$$

Algorithme : n : nombre de colonne de *dens* (=*T* dans notre cas)
m : nombre de lignes de *dens* (=2 dans notre cas)
ksi : matrice des probabilités prévisionnelles (*forecasted probabilities*) à m lignes et n+1 colonnes (2 x T+1 dans notre cas),

*ksi*1 : matrice des probabilités filtrés (*filtred probabilities*) à *m* lignes et *n* colonnes (2 x *T* dans notre cas)

$ksi(:,1) = Y_0$: la première colonne de *ksi* sera Y_0

Pour $indx = 1$ *jusqu'à n*

$zeta = ksi(:,indx).* dens(:,indx)$ Densité pondérée

$zeta = zeta / somme(zeta)$ Proba Filtrée

$ksi1(:,indx) = zeta$

$zeta1 = cp'* zeta$

$ksi(:,indx+1) = zeta1$ Proba Prévisionnelle

Boucle terminée

$$likk = \sum_{t}^{T} log \left| \sum_{j} ksi(j) \times dens(j) \right|$$

Outputs : *likk* : vraisemblance à maximiser

Calibrage des paramètres du modèle RSLN

Inputs : *par* : vecteur des paramètres initiés (6 paramètres dans le cas d'un modèle à 2 Régimes : $m_1, m_2, s_1, s_2, p_{12}, p_{21}$)

Algorithme : $\underset{par}{Max}(Likk) \backslash Y$: Maximiser *Likk* en *par* sachant les données observées

Outputs : *par* : les paramètres du modèle RSLN calibrés

Cette approche est mise en place grâce aux fonctions d'optimisation du logiciel MATLAB.

II-4-2 Données

Le calibrage est relatif à un modèle de mesure des rendements des actions (avec dividendes réinvestis). Cette mesure à la date *t* est égale à $\alpha_t = log\left(\frac{VIA_t}{VIA_{t-1}}\right)$, où VIA_t représente la valeur d'un investissement en action (disponible en lecture directe dans les données de Friggit [2007]).

II-4-3 Estimation des paramètres

Les estimations menées avec les données de Friggit [2007] sont comparées à celles obtenues par Hardy [2001] (les estimations d'Ahlgrim et al. [2005] portant sur un périmètre différent, elles ne sont pas comparables à celles obtenues dans la présente étude). Aussi, le tableau suivant recense les différents paramètres nécessaires à l'alimentation de l'algorithme du modèle sur les actions, en précisant les valeurs obtenues par Hardy [2001] d'une part, et celles obtenues dans le cadre du présent calibrage d'autre part. Nous noterons que le calibrage d'Hardy [2001] présenté est réalisé à partir d'un historique de l'indice américain S&P 500 couvrant la période 1956-1999 (pour mémoire les données de Friggit [2007] retenues dans la présente étude portent sur la France sur la période 1955-2005).

Paramètre	Intitulé	HARDY	FRIGGIT
μ^{ss}_1	Moyenne (régime 1, faible volatilité)	0,012	**-0,1211**
σ^{ss}_1	Écart type (régime 1, faible volatilité)	0,035	**0,0611**
$p^{ss}_{1,2}$	Probabilité de passage du régime 1 à 2	0,037	**0,6479**
$p^{ss}_{2,1}$	Probabilité de passage du régime 2 à 1	0,210	**0,1572**
μ^{ss}_2	Moyenne (régime 2, forte volatilité)	-0,016	**0,1661**
σ^{ss}_2	Écart type (régime 2, forte volatilité)	0,078	**0,1399**

Tab. 13 : *Estimation des paramètres du modèle RSLN pour les actions*

II-5 Calibrage du modèle de l'immobilier

Après avoir exposé la méthode retenue pour le calibrage ainsi que les données correspondantes, nous présentons les résultats obtenus des estimations des différents paramètres. Un test d'adéquation est finalement effectué.

II-5-1 Méthode

Pour mémoire, le modèle retenu pour l'immobilier est le suivant :

$$d\left(re\right)_t = \kappa_{re}\left(\mu_{re} - \left(re\right)_t\right)dt + \sigma_{re}dB_{re,t},$$

soit après discrétisation exacte :

$$\left(re\right)_{t+1} = \left(re\right)_t \times e^{\left(-\kappa_{re}\right)} + \mu_{re}\left(1 - e^{\left(-\kappa_{re}\right)}\right) + \varepsilon_{re,t} \times \sigma_{re} \times \sqrt{\frac{1 - e^{\left(-2\kappa_{re}\right)}}{2\kappa_{re}}}$$

II-5-2 Données

Cette sous-section décrit le calibrage d'un modèle sur une mesure des rendements de l'immobilier. Cette mesure à la date t est égale à $\left(re\right)_t = log\left(\frac{IPL_t}{IPL_{t-1}}\right)$, où IPL_t représente l'indice des prix des logements (disponible en lecture directe dans les données de Friggit [2007]).

II-5-3 Estimation des paramètres

Dans une première étape, l'estimation des paramètres est réalisée à partir d'un modèle autorégressif d'ordre 1 de la forme :

$$\left(re\right)_{t+1} = \alpha_{re} + \beta_{re} \times \left(re\right)_t + \varepsilon'_{re,t}$$

Nous obtenons les résultats suivants : $\hat{\alpha}_{re} = 0{,}01$ et $\hat{\beta}_{re} = 0{,}86$. Pour revenir au modèle initial décrit par Ahlgrim et al. [2005], les coefficients estimés sont alors transformés (**à l'image des**

travaux effectués dans le cadre du modèle d'inflation). Il en ressort les résultats suivants : $\hat{k}_{re} = 0{,}15$, $\hat{\mu}_{re} = 0{,}09$ et $\hat{\sigma}_{re} = 0{,}04$.

Le tableau suivant recense les différents paramètres nécessaires à l'alimentation de l'algorithme du modèle sur les rendements de l'immobilier, en précisant les valeurs obtenues par Ahlgrim et al. [2005] d'une part, et celles obtenues dans le cadre du présent calibrage d'autre part. Nous noterons que le calibrage d'Ahlgrim et al. [2005] est réalisé à partir de données américaines sur la période 1978-2001 (pour mémoire les données de Friggit [2007] retenues dans la présente étude portent sur la France sur la période 1955-2005).

Paramètre	**Intitulé**	**AHLGRIM ET AL.**	**FRIGGIT**
κ_{re}	Vitesse retour à la moyenne	1,2	**0,15**
μ_{re}	Rendement moyen de l'immobilier	0,023	**0,09**
σ_{re}	Écart type de l'erreur	0,013	**0,04**

Tab. 14 : *Estimation des paramètres du modèle de l'immobilier*

Les résultats du calibrage obtenus lors de l'étude avec les données de Friggit [2007] présente des différences significatives avec ceux obtenus par Ahlgrim et al. [2005].

Ces différences sont toutefois justifiées par la nature des données retenues par Ahlgrim et al. [2005], différente de celle de la présente étude. En effet, Ahlgrim et al. [2005] retiennent des données américaines trimestrielles sur la période 1978-2001, contre des données françaises annuelles sur la période 1955-2005 dans la présente étude. Aussi, Ahlgrim et al. [2005] ne tiennent pas compte des périodes 1955-1965 et 2001-2005 alors qu'elles correspondent en France à des niveaux de rendement élevés (en France, entre 1955-1965 le rendement annuel moyen est près de 20 % et entre 2001 et 2005 le rendement annuel moyen est passé de 7 % à 14 %). Ces niveaux de rendement élevés en France sur ces périodes augmentent la moyenne, la vitesse de retour à la moyenne et la volatilité des rendements.

Nous noterons que les paramètres présentés ne sont pas les paramètres estimés mais sont issus d'une transformation non linéaire de ceux-ci. Cette approche introduit un biais dans les paramètres et invite à apprécier la précision des estimateurs avec prudence.

II-5-4 Tests d'adéquation du modèle

Le tableau 15 reprend les principaux résultats sur les tests d'adéquation réalisés. Hormis le test sur la significativité de la constante, l'ensemble de ces tests valide les conditions posées *a priori* pour la validation du modèle et du calibrage (y compris l'hypothèse sur la normalité des résidus). Toutefois il est à noter que la significativité de la constante n'étant pas vérifiée, il conviendra d'interpréter le rendement moyen de l'immobilier avec prudence (le rendement moyen étant calculé à partir de la valeur de cette constante).

Test	Résultat	Interprétation
R^2	0,76	Le pouvoir explicatif du modèle est de 76 %
Fisher	$P_value = 2,2.10^{-16}$	Le modèle global est significatif
Student (constante)	$P_value = 0,12$	La constante n'est pas significative
Student (variable)	$P_value = 2,2.10^{-16}$	Le coefficient de la variable est donc significatif
Moyenne résidus	$1,5.10^{-18}$	La moyenne des résidus est considérée nulle
Durbin-Watson	$P_value = 0,16$	Absence d'auto-corrélation d'ordre 1
Breusch-Godfrey	$P_value = 0,29$	Absence d'auto-corrélation d'ordre $p = 1$
ARCH	$P_value = 0,93$	Absence d'hétéroscédasticité
Jarque-Bera	$P_value = 0,51$	Hypothèse de normalité des résidus vérifiée

Tab. 15 : *Résultats des tests d'adéquation pour le modèle de l'immobilier*

II-6 Calibrage du modèle de crédit

Dans ce qui suit, le calibrage est effectué pour les deux approches exposées dans le paragraphe I-2-2-3 de ce chapitre : l'approche des coefficients d'abattement et l'approche « mélange ».

- Approche des coefficients d'abattement

Le taux de marché *corporate* retenu, à titre illustratif, est égal[11] à 3,7 % (pour une maturité moyenne de 5 ans). Les obligations *corporate* utilisées lors de l'estimation de cette valeur sont présentées dans le tableau 16.

En supposant un taux nominal, pour la même maturité, de 3,2 % le coefficient d'abattement peut être obtenu de la façon suivante :

$$coeff_abtt = P_crédit / P_nominal = \frac{\frac{1}{(1+3,7\%)^5}}{\frac{1}{(1+3,2\%)^5}} = 97,6\%$$

[11] Cette valeur est estimée à partir d'une moyenne du taux actuariel d'un ensemble d'obligations *corporate* issu du lien suivant : http://www.financefi.fr/obligations/tableau.asp?donnees=corporate&pays=1&duree=3-6.

Nom de l'obligation	Durée de vie restante (ans)	TAUX ACTUARIEL (%)
CASA 4,10 % 10/04/14 (AA)*	4,912	3,83 %
CASINO 5,5% 31/01/2015 (BBB)*	5,17	4,37 %
CRED AGRI 4,90% 19/12/14 (AA)*	5,052	4,93 %
DANONE 5,5% 06/05/2015 (A)**	5,43	3,21 %
EDF 5,125% 23/01/2015 (AA-)*	5,148	3,11 %
SANOFI 3,125 % 10/10/2014 (AA-)*	4,86	2,71 %
SCHNEIDER 5,375% 08/01/2015 (A-)*	5,107	3,34 %
CARREFOUR 5,125% 10/10/2014 (BBB-)*	4,86	3,17 %
BOUYGUES 4,375% 29/10/2014 (BBB+)*	4,912	3,39 %
ARCELOR FIN 4,625% 07/11/2014 (BBB)***	4,937	4,00 %
ARCELOR FIN 5,5% 15/07/2014 (BBB)***	4,937	5,37 %

*Rating de Fitch IBCA
**Rating de Standard & Poor's
***Rating de Moody's

Tab. 16 : *Exemple de données utilisées pour l'estimation des paramètres du modèle crédit*

- Approche « mélange »

Pour rappel, l'approche « mélange » consiste à modéliser le *spread* à partir d'un modèle à changement de régime (modèle RSLN2).

Nous présentons ci-après des paramètres illustratifs estimés en se basant sur le même principe de calibrage que celui utilisé dans le cas des rendements des actions :

Spread	**Paramètres ajustés**
Valeur initiale	0,90 %
Matrice des probabilités de passage entre les deux régimes [[p_{11} p_{12}] ; [p_{21} p_{22}]]	[[70 % 30 %] ; [90 % 10 %]]
Moyenne du spread pour le régime 1 (faible volatilité)	0,60 %
Volatilité du spread pour le régime 1 (faible volatilité)	0,01 %
Moyenne du spread pour le régime 2 (forte volatilité)	0,90 %
Volatilité du spread pour le régime 2 (forte volatilité)	0,02 %

Tab. 17 : *Résultats d'estimation des paramètres du modèle crédit avec changement de régime*

III- Projection des scénarios

La présente section est relative à l'étape du *backtesting*. Plus précisément, elle présente et commente les résultats des projections du rendement, de la volatilité et des corrélations réalisées dans le cadre de la validation du modèle et de son paramétrage. Il est rappelé que le calibrage des modèles est utilisé à titre illustratif.

III-1 Éléments sur la mise en œuvre

Au niveau de la structure schématique de projection des scénarios, nous avons opté pour la structure linéaire (cf. sous-section II-1 du chapitre 1 de cette partie). Pour rappel, ci-dessous un exemple de cette structure lors d'une projection de 500 trajectoires de rendement annuel des actions entre 2010 et 2030 :

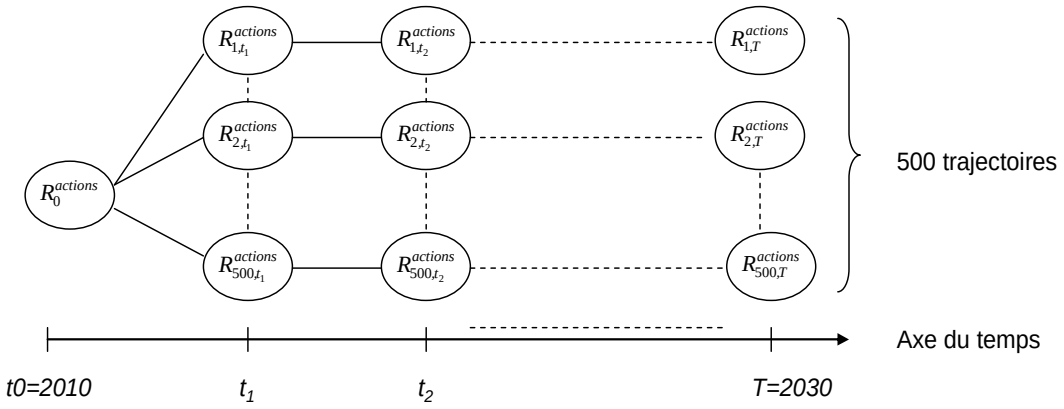


Fig. 15: *Illustration de la structure schématique linéaire dans le cas de la projection des rendements des actions*

En vue de simuler des browniens ayant une structure de corrélation prédéfinie, nous utilisons la décomposition de Choleski de leur matrice de variance covariance (cf. tableau 6).

En fait, si nous considérons un vecteur aléatoire gaussien $X = (X_1, X_2, ..., X_n)$ dont la distribution est multi-normale non dégénérée, sa densité est de la forme :

$$f_X(X) = \frac{1}{(2\pi)^{\frac{n}{2}} \times det(\Sigma)^{\frac{1}{2}}} \times exp\left[-\frac{1}{2}(x-\mu)'.\Sigma^{-1}.(x-\mu)\right]$$

Avec :

$\mu = (\mu_1, \mu_2, ..., \mu_n)$: Vecteur espérance des lois marginales

Σ : Matrice de variance-covariance

Plus précisément, cette dernière matrice s'écrit dans le cas de n=4:

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{pmatrix}$$

Comme Σ est symétrique et définie positive ($Y'\Sigma Y > 0, \forall Y \neq 0$), il est possible de trouver une matrice T unique, définie comme suit : $T = \begin{pmatrix} b_{11} & 0 & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & b_{33} & 0 \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix}$ tel que : $\Sigma = TT'$.

Cette dernière expression correspond à la décomposition de Choleski de Σ. Elle est utilisée ensuite pour la simulation de browniens qui reflètent la même structure de dépendance existante entre les composantes du vecteur X puisque nous savons que si Z est un vecteur indépendant identiquement distribué (i.i.d) de loi commune normale centrée réduite N (0,1), alors le vecteur $X = TZ + \mu$ a pour matrice de variance-covariance Σ et pour moyenne μ.

Nous utilisons cette relation pour retrouver de proche en proche les valeurs des coefficients de la matrice triangulaire T :

- $X_1 = b_{11}.Z_1 + \mu_1$

En considérant la variance de chacun des termes de cette égalité, nous obtenons b_{11}=1

- $X_2 = b_{21}.Z_1 + b_{22}.Z_2 + \mu_1$

En considérant la variance de chacun des termes de cette égalité, nous obtenons $b_{11} + b_{22}$=1
Si nous nous intéressons à la covariance, nous tombons sur l'égalité suivante :
$cov(X_1, X_2) = \rho_{12} = E[b_{11}.Z_1(b_{12}.Z_1 + b_{22}.Z_2)]$
Finalement, $b_{21} = \frac{\rho_{12}}{b_{11}} = \rho_{12}$ et $b_{21} = (1 - \rho_{12}^2)^{\frac{1}{2}}$

- Etc.

Le graphique suivant illustre le cas où nous souhaitons simuler des variables négativement corrélées. Cela passe par la génération dans un premier temps de browniens indépendants (composants le vecteur Z) i.i.d et de loi N (0,1).

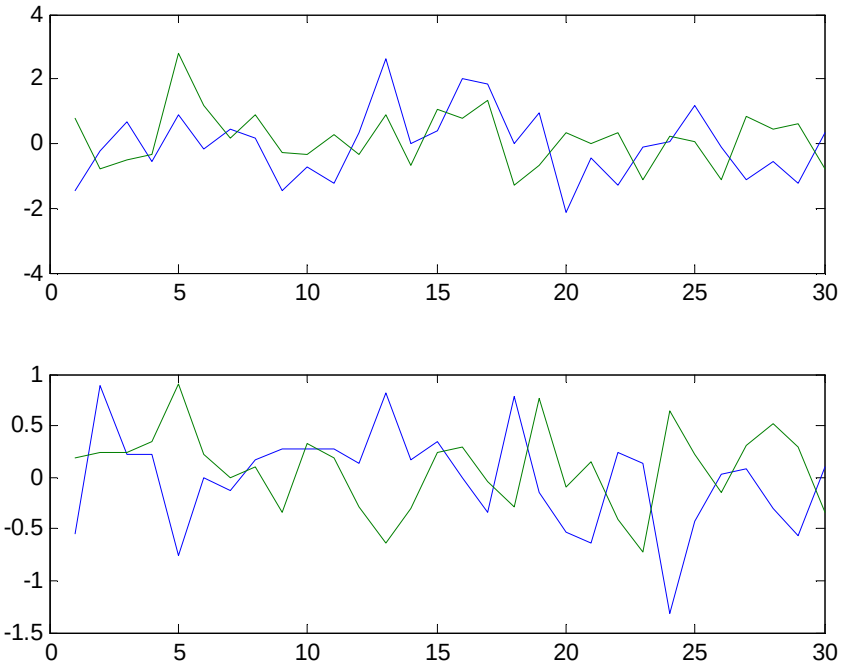


Fig. 16 : *Illustration du passage de browniens indépendants à des browniens corrélés négativement*

Notons que ce modèle suppose que les taux d'intérêts suivent la loi normale, donc il y a une possibilité d'obtenir des taux réels négatifs. Deux bornes, une borne supérieure et une autre inférieure, peuvent être fixée par l'investisseur pour les taux instantanés et les taux longs.

Les modèles développés (un modèle pour les taux courts, notés r_t, et un modèle pour les taux longs, notés l_t) reprennent l'approche de retour à la moyenne du modèle de Hull et White [1994] à deux facteurs :

$$dr_t = \kappa_r \left(l_t - r_t \right) dt + \sigma_r dB_{r,t}$$
$$dl_t = \kappa_l \left(\mu_l - l_t \right) dt + \sigma_l dB_{l,t}$$

Nous rappelons que la discrétisation retenue est la discrétisation exacte (cf. Planchet et al. [2005]) [12] :

$$r_{t+1} = r_t \times e^{(-\kappa_r)} + l_t \left(1 - e^{(-\kappa_r)}\right) + \varepsilon_{r,t} \times \sigma_r \times \sqrt{\frac{1 - e^{(-2\kappa_r)}}{2\kappa_r}}$$

$$l_{t+1} = l_t \times e^{(-\kappa_l)} + \mu_l \left(1 - e^{(-\kappa_l)}\right) + \varepsilon_{l,t} \times \sigma_l \times \sqrt{\frac{1 - e^{(-2\kappa_l)}}{2\kappa_l}}$$

[12] La discrétisation exacte doit être utilisée lorsqu'elle est disponible, cette discrétisation étant la seule à ne pas introduire de biais sur la loi du facteur discrétisé.

où :

κ_r (resp. κ_l) est la vitesse de retour à la moyenne ;

μ_l est le taux d'intérêt à long terme moyen.

D'un autre côté et compte tenu de l'absence de données fiables pour la partie long terme, il a été retenu pour la moyenne de long terme l'objectif BCE de 2 % et pour les deux autres paramètres des niveaux permettant d'avoir des résultats cohérents, sans aucun calibrage statistique.

Mise en œuvre de la démarche crédit

La mise en œuvre de la démarche repose sur la spécification du prix des obligations sans risque de défaut et sur le prix observé sur le marché des obligations présentant un risque de défaut.

Les prix $O(T)$ des obligations sans risque de défaut de nominal *N*, de maturité *T* et de taux facial γ se déduisent via la formule suivante (cf. Planchet et al. [2005]) :

$$O(T) = N \times \left(\gamma \sum_{i=1}^{T} P_l(0,i) + P_l(0,T) \right)$$

où $P_l(t,T)$ représente le prix à la date *t* des zéro-coupon de maturité *T*. La démarche présentée consiste à déterminer l'abattement α tel que :

$$S(T) = \alpha \times O(T)$$

où $S(T)$ représente le prix de marché observé pour les obligations de maturité *T* avec risque de défaut. En pratique, nous noterons que α dépend de la maturité et du rating des obligations présentant un risque de défaut.

Ci-après des paramètres relatifs à l'environnement du calcul. Notons qu'avec 2 000 simulations nous obtiendrons des résultats qui sont égaux aux résultats observés avec un nombre de simulations supérieur.

Horizon de projection (en années)	50
Nombre de pas	200
Nombre de simulations	2 000

Tab. 17 : *Paramètres utilisés dans le GSE construit*

III-2 Tests du modèle : paramètres et matrice de corrélation

Cette sous-section est relative aux tests du modèle, et en particulier aux tests sur les paramètres et la matrice de corrélation. L'analyse et l'interprétation de ces tests devra être traitée avec prudence (existence de biais dus à l'absence de linéarité dans les transformations effectuées, prise en compte de la périodicité des projections, etc.).

Face aux limites de l'estimation des paramètres des modèles de GSE, en particulier à cause de la non évidence de l'hypothèse de la reproduction de l'historique dans le futur, nous supposerons que le seul garant de la fiabilité du GSE considéré par la suite sera le degré de correspondance des résultats obtenus (projections) par rapport à ceux prévus lors du calibrage (avis des experts, paramètres historiques, anticipations à partir des valeurs actuelles de marché des différentes produits et variables financières, etc.). Le respect de cette correspondance sera considéré comme une condition minimale à vérifier.

III-2-1 Tests sur les paramètres et sur les développements informatiques

Les tests sur les paramètres peuvent être organisés en deux parties.

› Comparaison de valeurs théoriques et de valeurs empiriques

Le premier point est relatif à la comparaison de valeurs théoriques et de valeurs empiriques estimées sur les trajectoires simulées. À cet effet, il convient de distinguer différentes classes d'actifs.

Cas des actions et de l'immobilier

Pour les actions et l'immobilier, nous pouvons comparer les valeurs théoriques et empiriques des rendements ou des volatilités. Soit $x_{i,t}, x_{i,t+1}, ..., x_{i,t+h}$ la trajectoire des rendements empiriques de la simulation *i* jusqu'à l'horizon $t+h$:

- le rendement moyen empirique est égal à $\hat{\mu} = \frac{1}{I} \times \frac{1}{H+1} \sum_{i=1}^{I} \sum_{h=0}^{H} x_{i,t+h}$ (où I représente le nombre total de simulations et H l'horizon de la projection) ;

- la volatilité moyenne empirique est calculée par $\hat{\sigma} = \sqrt{\frac{1}{I+H+1} \times \sum_{i,h}^{I,H} \left(x_{i,t+h} - \hat{\mu} \right)^2}$.

Cas des produits monétaires

Dans le cas des produits monétaires, la comparaison présentée ci-dessus peut également être appliquée.

Cas des produits obligataires

Concernant les produits obligataires, la comparaison en lecture directe est plus difficile car dans les sorties du modèle nous disposons du prix des obligations, desquels sont déduits les rendements et les volatilités associés, alors que les paramètres retenus pour ces classes d'actifs portent sur la moyenne des taux d'intérêt réel et des taux d'inflation.

La comparaison des valeurs théoriques et empiriques pourrait toutefois être effectuée à partir des formules de prix des obligations zéro-coupon, sans tenir compte du risque de défaut. En effet, par exemple dans les modèles de type Vasicek à un-facteur les prix à la date *t* des zéro-coupon de maturité *T*, notés $P_l(t,T)$, sont déterminés à partir des taux *r* (κ représentant la vitesse de retour à la moyenne et μ représentant le taux moyen) par la relation suivante (cf. Hull [2000]) :

$$P(t,T) = A(t,T)e^{(-B(t,T)\times r)}$$

Avec :

$$B(t,T) = \frac{1-e^{(-\kappa(T-t))}}{\kappa}$$

$$A(t,T) = exp\left(\frac{(B(t,T)-T+t)(\kappa^2\times\mu-\sigma^2/2)}{\kappa^2} - \frac{\sigma^2 B(t,T)^2}{4\kappa}\right)$$ [13]

Il convient ainsi de comparer le prix théorique $P_l(t,T)$ obtenu avec le paramètre de taux moyen au prix empirique $\hat{P}_l(t,T)$ obtenu avec la moyenne (sur l'ensemble des simulations et l'ensemble de la durée de projection) des taux projetés.

› Tests aux limites (cas du modèle des actions)

Le deuxième point est relatif aux tests aux limites. À cet effet, il s'agit par exemple de retenir un seul régime pour le RSLN2 en mettant une probabilité de transition à 0 ou de considérer des versions déterministes en mettant les volatilités à 0. Nous pouvons également vérifier pour le modèle de taux que nous retrouvons le modèle à un facteur en mettant les paramètres du taux long à 0. Dans le même esprit, nous pouvons vérifier que les sens de variation des grandeurs modélisées sont conformes à la théorie (observer ce qui se passe lorsque la volatilité augmente sur une option, une obligation, etc.).

III-2-2 Tests sur les matrices de corrélation

Une fois les tests sur les paramètres et les développements informatiques réalisés, il convient de décrire un test de cohérence du paramétrage consistant à estimer empiriquement la matrice de corrélation des rendements et à la comparer à celle que nous observons sur le marché (si elle est disponible). En fait la matrice de corrélation théorique (*input* du GSE) et par la suite la matrice de corrélation empirique (*output* du GSE) peut différer de celle observée réellement sur le marché (avec des données historiques) en fonction des prévisions de l'investisseur. Cela ne remet pas en cause l'intérêt de la comparaison entre ces deux dernières matrices par exemple pour effectuer des tests de sensibilité.

La matrice que nous observons sur le marché prend également la forme suivante :

$$\begin{pmatrix} \rho_{1,1} & \rho_{1,2} & \cdots & \cdots & \rho_{1,K} \\ \rho_{2,1} & \rho_{2,2} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \rho_{K,1} & \cdots & \cdots & \cdots & \rho_{K,K} \end{pmatrix}$$

où $\rho_{i,j}$ représente la corrélation entre les classes d'actifs numéros k_1 et k_2, et est déterminée par :

[13] *exp* désigne ici la fonction exponentielle.

$$\rho_{k_1,k_2} = \frac{\mathbf{cov}(k_1,k_2)}{\sigma(k_1)\times\sigma(k_2)} = \frac{\sum_{i,h}^{I,H}\left(\left(x_{k_1}\right)_{i,h} - \overline{\left(x_{k_1}\right)}\right)\left(\left(x_{k_2}\right)_{i,h} - \overline{\left(x_{k_2}\right)}\right)}{\sqrt{\sum_{i,h}^{I,H}\left(\left(x_{k_1}\right)_{i,h} - \overline{\left(x_{k_1}\right)}\right)^2}\cdot\sqrt{\sum_{i,h}^{I,H}\left(\left(x_{k_2}\right)_{i,h} - \overline{\left(x_{k_2}\right)}\right)^2}}$$

avec :

$\left(x_{k_1}\right)$ représentant le rendement des actifs de la classe k_1 ;

$\left(x_{k_2}\right)$ représentant le rendement des actifs de la classe k_2 ;

i représentant le nombre de simulations (*I* simulations au total) ;

h représentant l'année de projection (*H* années de projection au total).

III-3 Résultats de projection

Nous présentons ci-après des graphiques mettant en évidence quelques propriétés théoriques du modèle de GSE objet de l'étude (exemple : retour à la moyenne pour les taux, changement de régime pour les actions, etc.) ainsi que certains résultats sur les niveaux des rendements.

Rendements moyens et volatilités sur toute la période de projection

Le tableau suivant illustre la volatilité et le rendement moyen obtenus sur toute la période de projection. En particulier, nous observons que la volatilité est cohérente avec le rendement.

Pour le monétaire, les actions et l'immobilier nous observons que les rendements moyens et les volatilités sont du même ordre de grandeur que les paramètres du calibrage ce qui est donc cohérent avec les approches retenues pour modéliser ces actifs.

Rendements moyens et volatilités	Rendement	Volatilité
Monétaire	3,51 %	1,52 %
ZC 5 ans (nominal)	4,25 %	4,42 %
ZC 10 ans (nominal)	4,32 %	6,67 %
ZC 15 ans (nominal)	4,34 %	7,46 %
ZC 8 ans (réel)	3,59 %	3,96 %
ZC 15 ans (réel)	3,61 %	4,49 %
Crédit 5 ans	4,27 %	4,42 %
Actions	7,22 %	16,53 %
Immobilier	5,91 %	8,81 %

Tab. 18 : *Exemple des projections de rendement obtenues par le GSE construit*

La prime de risque entre les obligations nominales et les obligations indexées est égale à près de 75 points de base pour les ZC 15 ans et apparaît *a priori* importante. Nous noterons toutefois qu'entre 2004 et 2009 certains fonds présentent des primes de risques pour ces

classes d'actifs de près de 100 points de base[14]. Sauf précision contraire, tous les chiffres présentés ci-après sont établis à partir des résultats ci-dessus obtenus dans le cadre du scénario central.

Courbe des taux des obligations indexées sur l'inflation d'une maturité de 15 ans

La courbe des taux présentée ci-après en bleu (cf. graphique 17) correspond à la courbe des taux déduite des valeurs des taux de rendement annualisés et projetés de l'obligation indexées d'une maturité de 15 ans avec un pas trimestriel (ce qui correspond à un nombre de 60 pas au total). En rouge une approximation par une fonction linéaire par morceau est tracée. Cette fonction correspond aux quatre droites dont les pentes sont calculées respectivement sur les quatre périodes 0-2 ans, 2-5 ans, 5-10 ans, 10-15 ans. Il est précisé que l'unité de temps considéré dans le graphique est trimestrielle. Ces pentes ont les valeurs suivantes :

Périodes	Période 0-2	Période 2-5	Période 5-10	Période 10-15
Unité de temps trimestrielle	0,15 %	0,04 %	0,01 %	0,0046 %
Unité de temps annuelle	0,59 %	0,15 %	0,06 %	0,02 %

Tab. 19 : *Pentes de la courbe des taux obtenues via le GSE construit*

Rappelons que la pente sur une période s'étalant entre t et $t+h$ est la différence entre le taux à l'instant $t+h$ et le taux à l'instant t le tout divisé par h.

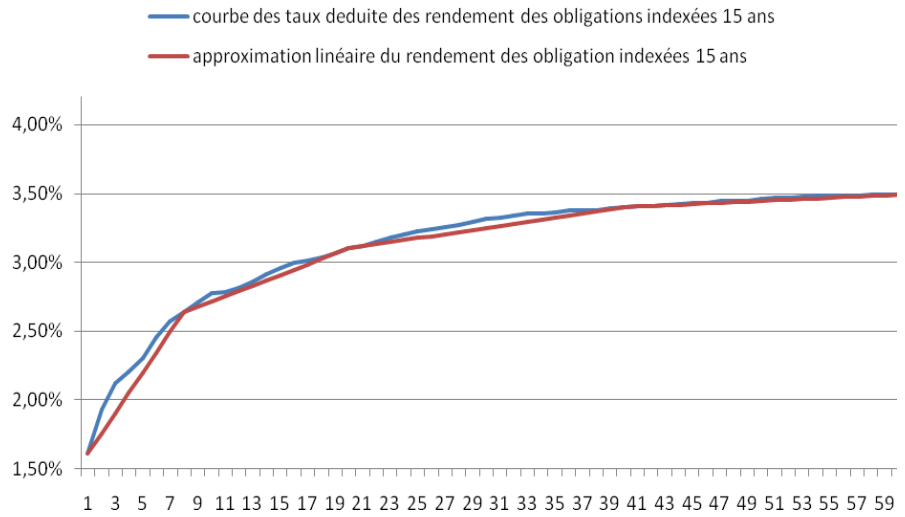


Fig.17 : *Courbe des taux des obligations indexées sur l'inflation d'une maturité de 15 ans*

[14] cf. par exemple les fonds suivants :
http://www.sgam.fr/portal/site/sgamfr/menuitem.8d6c6203bb02b5bf1398c6e5100000f7/?vgnextoid=5ce99634dd134010VgnVCM1000000100007fRCRD&countryiso=FR&spaceid=1&lang=fr&isin=LU0123756628

http://www.sgam.fr/portal/site/sgamfr/menuitem.8d6c6203bb02b5bf1398c6e5100000f7/?vgnextoid=5ce99634dd134010VgnVCM1000000100007fRCRD&countryiso=FR&spaceid=1&lang=fr&isin=LU0197574634.

Une autre façon de construire la courbe des taux sera de partir des taux courts simulés selon les modèles de taux proposés (Hull et White à 2 facteurs), pour ensuite déterminer les taux sur différents horizons (ou maturités) en utilisant la formule présentée dans le paragraphe I-2-2-1 du chapitre 2, à savoir : $R(t,T) = -log\{P(t,T)\}/(T-t)$.

Dans ce cadre, nous avons considéré le cas où le taux réel long observé initialement $l_r(0)$ est significativement plus élevé que le niveau moyen de long terme μ_{lr}. Cela donne lieu à une courbe des taux réels avec une inflexion. La STTI des taux nominaux aura également cette même forme puisque elle est déduite à la fois à partir des taux réels et des taux d'inflation (anticipés). Nous confirmons ainsi l'avantage, précédemment cité, du choix d'un modèle de taux à deux facteurs : ce dernier permet d'obtenir différentes formes de la courbe des taux en conformité avec ce qui est observé réellement (cf. Martellini et al. [2003]).

Le phénomène de retour à la moyenne, qui par hypothèse concerne les variables de taux du GSE construit, peut être mis en évidence au niveau des courbes de taux elles-mêmes. Le graphique 19 illustre ce cas de figure. En construisant les STTI à six dates différentes, à savoir (la date initiale et cinq dates futures : fin de la deuxième année, fin de la cinquième année, fin de la dixième année, fin de la quinzième année et fin de la vingtième année), nous constatons à partir d'une certaine date une certaine stabilité de la courbe La maturité maximale représentée par la STTI est ici 30 ans.

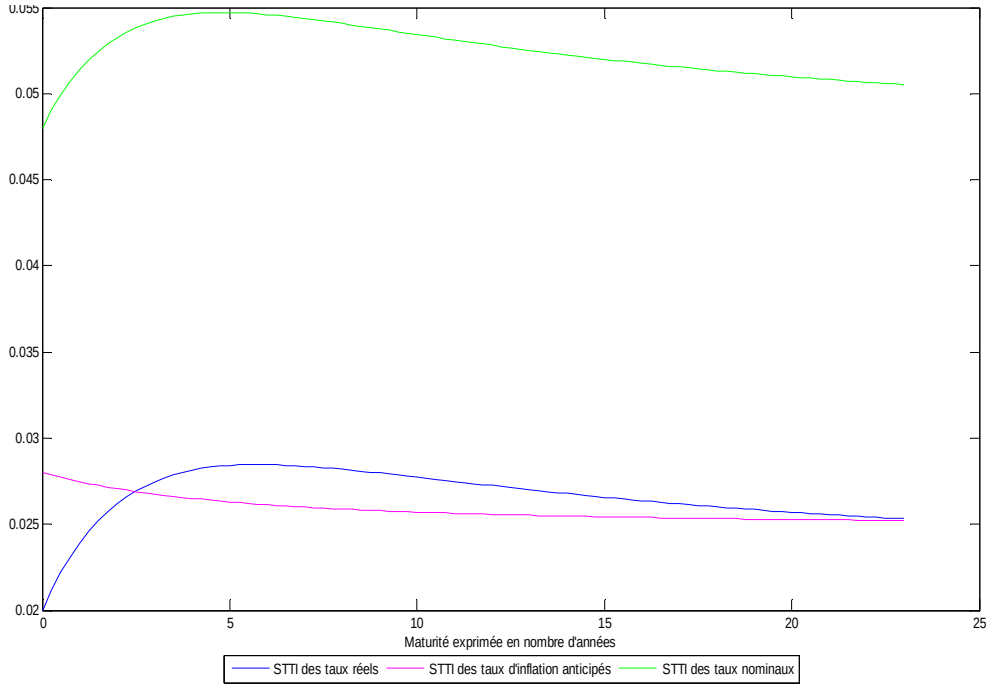


Fig. 18 : *Structure par terme des taux nominaux, des taux réels et des taux d'inflation en cas de taux réels longs élevés par rapport à ceux de l'équilibre : Mise en évidence de la courbure de la STTI*

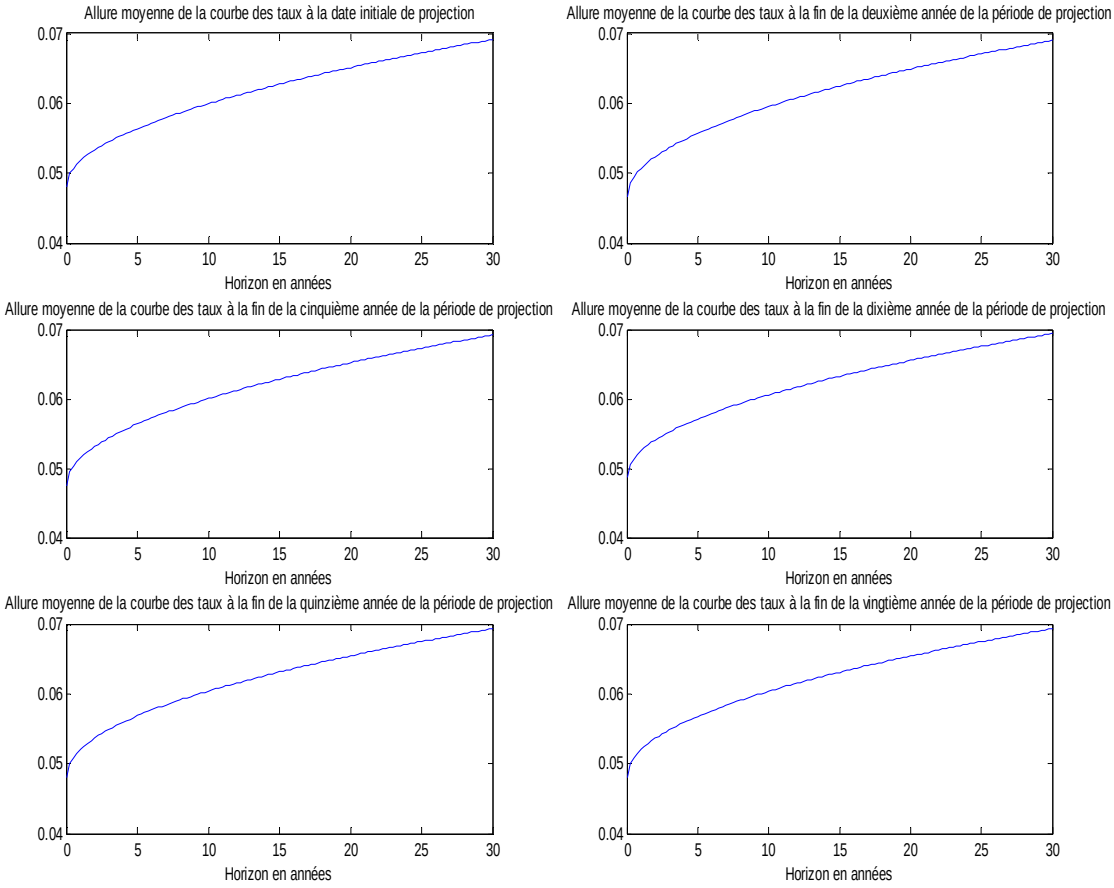


Fig.19 : *Evolution dans le temps de la STTI : Mise en évidence du phénomène de retour à la moyenne*

Ce même phénomène est confirmé pour ce qui concerne les écarts absolus moyens, cette fois-ci, entre les taux de deux maturités différentes (cf. graphique 20). Nous appellerons « pentes moyennes de taux » les écarts absolus moyens entre deux taux de maturités différentes évalués à la même date, ils sont calculés respectivement sur les quatre périodes 0-2 ans, 2-5 ans, 5-10 ans, 10-15 ans. Ces écarts, exprimés en point de base, se stabilisent après cinq années autour de leurs niveaux moyens. De même, nous constatons une décroissance de cet écart en augmentant les maturités considérées (passage d'un écart moyen de 57 points de base (pb) entre les taux de maturité 2 ans et ceux instantanés à 28 pb entre les taux de maturité 15 ans et 10 ans). Cela est cohérent avec les hypothèses considérées d'une faible volatilité sur les niveaux des taux longs.

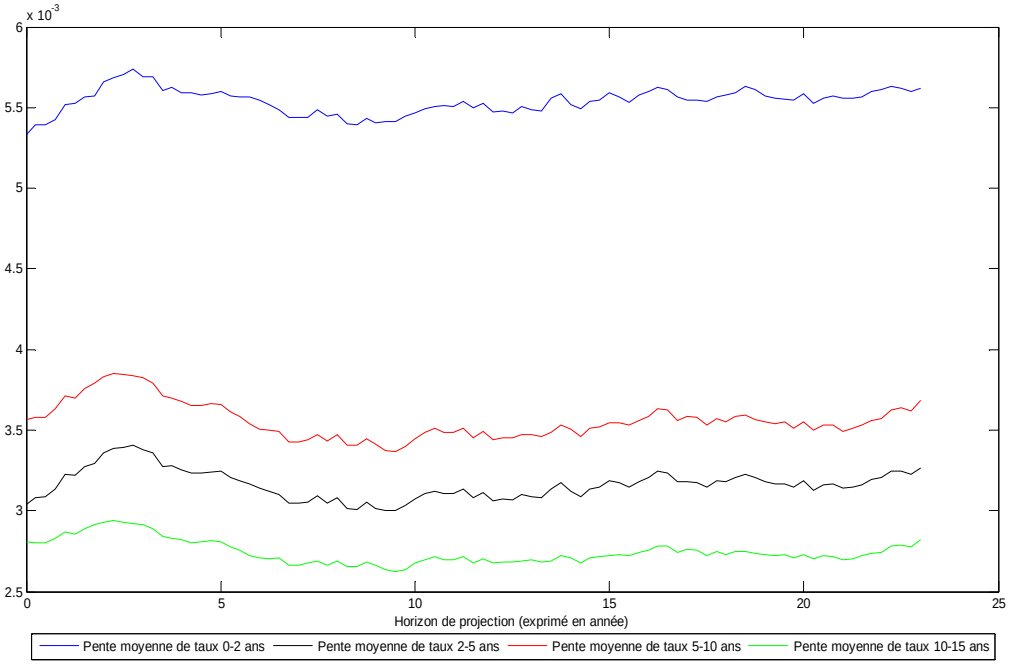


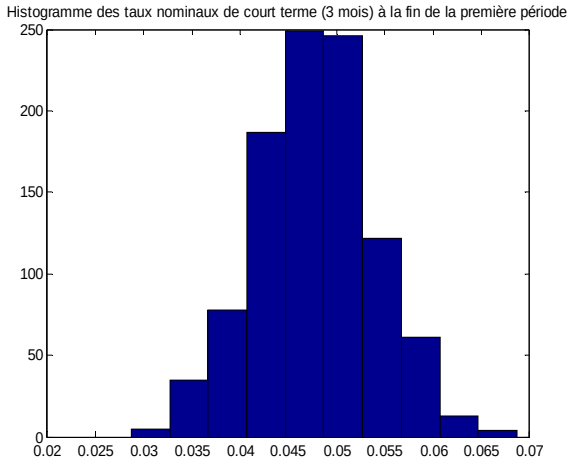
Fig. 20 : *Evolution des pentes moyennes de taux (écarts absolus moyens entre les taux de deux maturités différentes) tout au long de la période de projection*

	Pente 0-2 ans	Pente 2-5 ans	Pente 5-10 ans	Pente 10-15 ans
Moyennne projetée*	57	33	37	28

* Moyenne exprimée en point de base (et sur toute la période de projection)

Tab. 20 : *Moyennes des différentes pentes de taux projetées*

Les graphiques ci-dessous représentent, sous forme d'histogrammes, les distributions obtenues de différentes variables de taux et/ou de rendement. Les histogrammes permettent d'avoir une vision globale sur la fréquence des niveaux de taux et/ou de rendement simulés. Certes, il existe plusieurs tests statistiques pour juger de la conformité des valeurs projetées avec les hypothèses de départ retenues dans le modèle (normalité, etc.). Cependant l'étude des modèles pris individuellement ne constitue pas notre principal centre d'intérêt et c'est plutôt la structuration de ces différents modèles entre eux, dans l'objectif de la construction d'un GSE, qui suscite le plus notre attention.



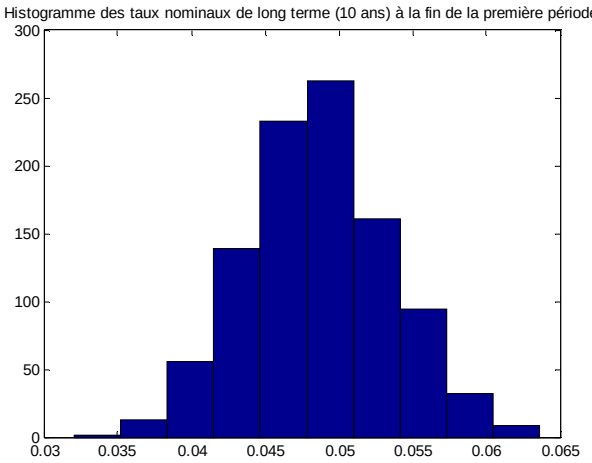
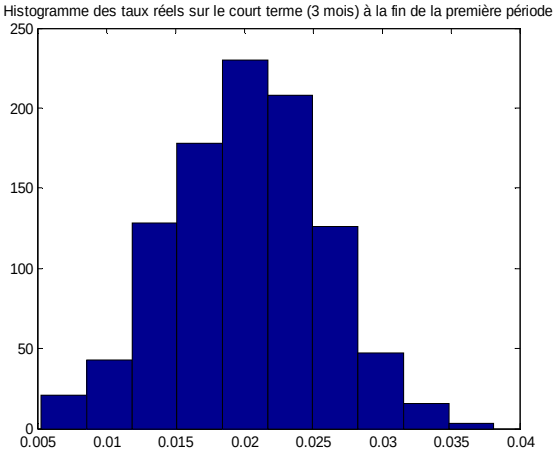


Fig. 21 : *Histogramme des taux nominaux de court terme (3 mois) à la fin de la première période*
Fig. 22 : *Histogramme des taux nominaux de long terme (10 ans) à la fin de la première période*



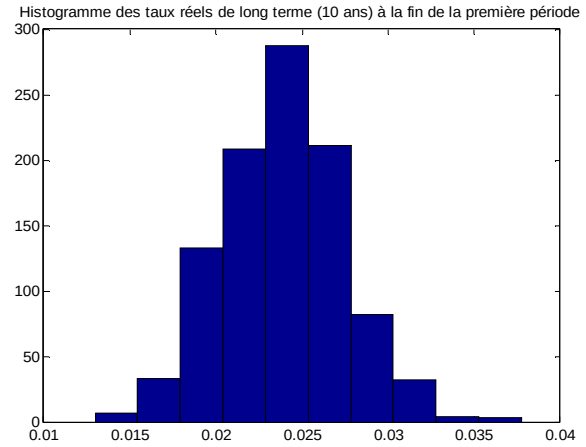
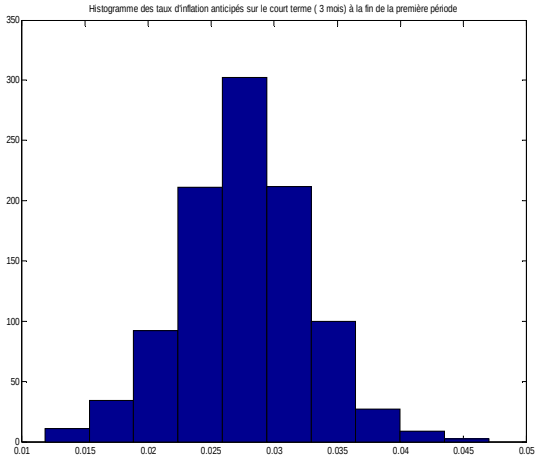


Fig. 23 : *Histogramme des taux réels sur le court terme (3 mois) à la fin de la première période*
Fig. 24 : *Histogramme des taux réels de long terme (10 ans) à la fin de la première période*



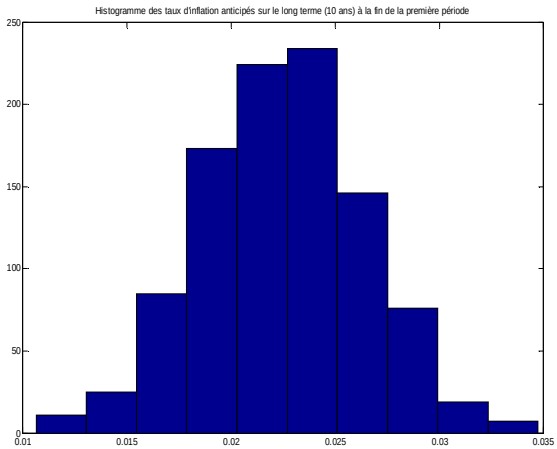
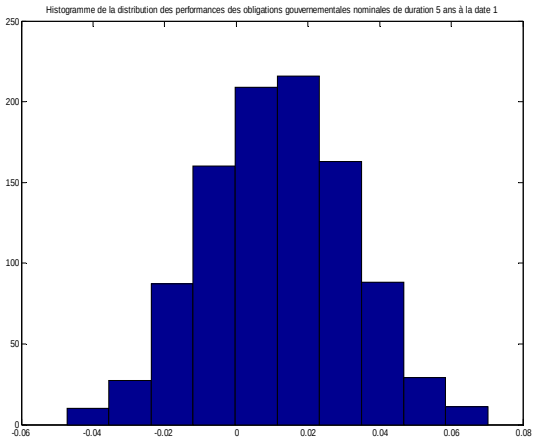


Fig. 25 : *Histogramme des taux d'inflation anticpés de court terme (3 mois) à la fin de la première période*
Fig. 26 : *Histogramme des taux d'inflation anticpés de long terme (10 ans) à la fin de la première période*



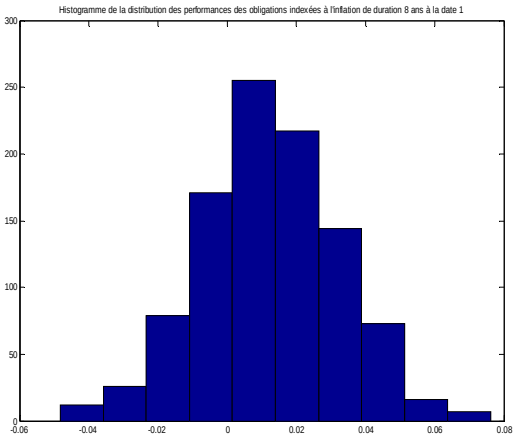


Fig. 27 : *Histogramme de la distribution des obligations nominales gouvernementales de duration 5 ans sur la première période*

Fig. 28 : *Histogramme de la distribution des obligations indexées inflation de duration 8 ans sur la première période*

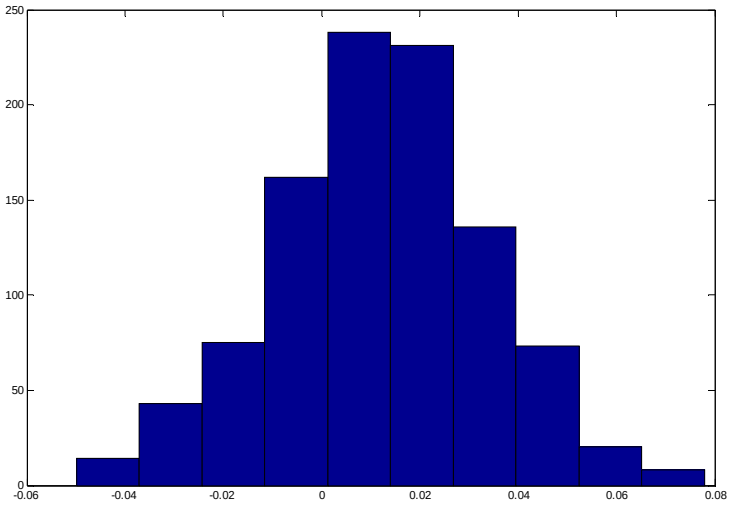


Fig. 29 : *Histogramme de la distribution des rendements des obligations nominales non gouvernementales de duration 5 ans sur la première période*

Matrice de corrélation des browniens sur la période de projection

Il est important de préciser que les valeurs retenues dans la matrice de corrélation des browniens, utilisée en *input* du GSE, se basent sur des ajustements : certains niveaux de corrélations sont en effet attribués de façon subjective sans nécessairement s'aligner sur les résultats historiques. Cela nous a conduits à proposer la matrice suivante à titre illustratif :

	Action 1	Inflation longue	Inflation	Taux long réel	Taux court réel	Monétaire	Immobilier	Action 2
Action 1	100,00 %	-20,00 %	0,00 %	-10,00 %	0,00 %	1,00 %	75,00 %	0,00 %
Inflation longue	-20,00 %	100,00 %	10,00 %	-20,00 %	0,00 %	0,00 %	-20,00 %	40,00 %
Inflation	0,00 %	10,00 %	100,00 %	0,00 %	-30,00 %	20,00 %	30,00 %	-40,00 %
Taux long réel	-10,00 %	-20,00 %	0,00 %	100,00 %	10,00 %	0,00 %	-20,00 %	30,00 %
Taux court réel	0,00 %	0,00 %	-30,00 %	10,00 %	100,00 %	-1,00 %	-40,00 %	40,00 %
Monétaire	1,00 %	0,00 %	20,00 %	0,00 %	-1,00 %	100,00 %	6,00 %	0,00 %
Immobilier	75,00 %	-20,00 %	30,00 %	-20,00 %	-40,00 %	6,00 %	100,00 %	-37,00 %
Action 2	0,00 %	40,00 %	-40,00 %	30,00 %	40,00 %	0,00 %	-37,00 %	100,00 %

Tab. 21 : *Corrélations théoriques entre les différentes variables (input du GSE)*

La matrice ci-après illustre les corrélations entre les browniens projetés (obtenues après projection de 1000 scénarios). Nous observons que les valeurs affichées sont très proches des valeurs présentées dans la matrice de corrélation des browniens en *input*. Ces valeurs correspondent aux valeurs attendues et sont donc cohérentes.

Corrélations entre les browniens (modèle BS 2 régimes)	Action 1	Inflation longue	Inflation	Taux long réel	Taux court réel	Monétaire	Immobilier	Action 2
Action 1	100,00 %	-19,58 %	-0,39 %	-10,06 %	-0,21 %	1,03 %	74,94 %	0,00%
Inflation longue	-19,58 %	100,00 %	9,97 %	-19,72 %	0,17 %	0,04 %	-19,81 %	40,10 %
Inflation	-0,39 %	9,97 %	100,00 %	0,37 %	-29,77 %	19,94 %	29,64 %	-39,98 %
Taux long réel	-10,06 %	-19,72 %	0,37 %	100,00 %	10,02 %	0,22 %	-19,94 %	29,84 %
Taux court réel	-0,21 %	0,17 %	-29,77 %	10,02 %	100,00 %	-0,94 %	-40,07 %	39,77 %
Monétaire	1,03 %	0,04 %	19,94 %	0,22 %	-0,94 %	100,00 %	5,95 %	0,09 %
Immobilier	74,94 %	-19,81 %	29,64 %	-19,94 %	-40,07 %	5,95 %	100,00 %	-36,95 %
Action 2	0,00 %	40,10 %	-39,98 %	29,84 %	39,77 %	0,09 %	-36,95 %	100,00 %

Tab. 22 : *Corrélations obtenues lors des projections entre les différentes variables*

Matrice de corrélation des rendements sur la période de projection

Nous nous intéressons ici à la dépendance entre eux des rendements des indices projetés. Cette dépendance est reflétée par les corrélations entre les browniens qui interviennent dans la dynamique des variables modélisées. Les valeurs cibles sont en pratique fournies sur la base d'avis d'experts, d'une étude historique ou encore déduits de manière implicite de l'observation du marché. Il s'agit en fait d'une matrice de corrélations jugée « raisonnable » par la place.

Globalement les valeurs obtenues dans le cadre du présent *backtesting* sont cohérentes avec les valeurs théoriques utilisées initialement en *input*. Nous notons que seules les corrélations entre l'immobilier et les obligations présentent des différences entre les valeurs cibles et celles obtenues par *backtesting*.

Plus particulièrement nous remarquons que :

- Les corrélations entre les zéro-coupons (ZC) nominaux et les ZC réels sont de l'ordre de 45 % à 55 % et donc sont cohérentes avec les valeurs cibles (pour mémoire ces valeurs sont de l'ordre de 60 %).

- La corrélation entre les actions et l'immobilier est de plus de 20 %, ce qui est cohérent avec la valeur cible (26 %).

- Les corrélations entre les ZC réels sont très fortes (de l'ordre de 99 %). Ces fortes corrélations sont d'un coté cohérentes avec l'approche du calcul des prix de ces obligations et représentent, d'un autre coté des niveaux proches de ceux de la cible. Elles traduisent le caractère très contraint de la courbe des taux, déterminée dans le cadre d'un modèle à deux facteurs.

Corrélations entre les indices projetés	Monétaire	ZC 5 ans (nominal)	ZC 10 ans (nominal)	ZC 15 ans (nominal)	ZC 8 ans (réel)	ZC 15 ans (réel)	Obligation crédit	Actions	Immobilier
Monétaire	100,00 %	-7,92 %	-8,79 %	-9,80 %	0,46 %	0,40 %	-7,92 %	1,06 %	1,45 %
ZC 5 ans (nominal)	-7,92 %	100,00 %	98,56 %	97,32 %	56,49 %	55,44 %	99,99 %	14,38 %	1,23 %
ZC 10 ans (nominal)	-8,79 %	98,56 %	100,00 %	99,73 %	46,91 %	46,80 %	98,56 %	15,85 %	1,45 %
ZC 15 ans (nominal)	-9,80 %	97,32 %	99,73 %	100,00 %	42,78 %	42,94 %	97,32 %	15,80 %	1,10 %
ZC 8 ans (réel)	0,46 %	56,49 %	46,91 %	42,78 %	100,00 %	99,68 %	56,49 %	5,28 %	5,66 %
ZC 15 ans (réel)	0,40 %	55,44 %	46,80 %	42,94 %	99,68 %	100,00 %	55,44 %	5,83 %	5,77 %
Obligation crédit	-7,92 %	99,99 %	98,56 %	97,32 %	56,49 %	55,44 %	100,00 %	14,38 %	1,23 %
Actions	1,06 %	14,38 %	15,85 %	15,80 %	5,28 %	5,83 %	14,38 %	100,00 %	20,72 %
Immobilier	1,45 %	1,23 %	1,45 %	1,10 %	5,66 %	5,77 %	1,23 %	20,72 %	100,00 %

Tab. 23 : *Corrélations obtenues lors des projections entre les différents rendements*

Corrélation des actifs dans un régime à faible volatilité

La matrice suivante présente la matrice de corrélation des actifs lorsque les actions évoluent dans un régime à faible volatilité. Nous remarquons que dans ce régime l'immobilier par exemple est positivement corrélé avec les actions.

Corrélation régime 1 (faible volatilité)	Monétaire	ZC 5 ans (nominal)	ZC 10 ans (nominal)	ZC 15 ans (nominal)	ZC 8 ans (réel)	ZC 15 ans (réel)	Obligation crédit	Actions	immobilier
Monétaire	100,00 %	-7,92 %	-8,77 %	-9,78 %	0,72 %	0,71 %	-7,92 %	1,03 %	1,62 %
ZC 5 ans (nominal)	-7,92 %	100,00 %	98,56 %	97,30 %	56,35 %	55,29 %	100,00 %	15,15 %	1,23 %
ZC 10 ans (nominal)	-8,77 %	98,56 %	100,00 %	99,73 %	46,70 %	46,57 %	98,56 %	16,63 %	1,54 %
ZC 15 ans (nominal)	-9,78 %	97,30 %	99,73 %	100,00 %	42,53_%	42,68 %	97,30 %	16,50 %	1,23 %
ZC 8 ans (réel)	0,72 %	56,35 %	46,70 %	42,53 %	100,00 %	99,68 %	56,35 %	5,92 %	5,37 %
ZC 15 ans (réel)	0,71 %	55,29 %	46,57 %	42,68 %	99,68 %	100,00 %	55,29 %	6,49 %	5,49 %
Obligation crédit	-7,92 %	100,00 %	98,56 %	97,30 %	56,35 %	55,29 %	100,00 %	15,15 %	1,23 %
Actions	1,03 %	15,15 %	16,63 %	16,50 %	5,92 %	6,49 %	15,15 %	100,00 %	21,88 %
Immobilier	1,62 %	1,23 %	1,54 %	1,23 %	5,37 %	5,49 %	1,23 %	21,88 %	100,00 %

Tab. 24 : *Corrélations obtenues lors des projections entre les différents rendements dans un régime à faible volatilité*

Corrélation des actifs dans un régime à forte volatilité

La matrice suivante présente la matrice de corrélation des actifs lorsque les actions évoluent dans un régime à forte volatilité. Nous remarquons que dans ce régime l'immobilier est négativement corrélé avec les actions. Cela peut s'expliquer par l'hypothèse de départ considérée : l'immobilier est considéré comme une valeur refuge en cas de crise.

Corrélation en régime 2 (forte volatilité)	Monétaire	ZC 5 ans (nominal)	ZC 10 ans (nominal)	ZC 15 ans (nominal)	ZC 8 ans (réel)	ZC 15 ans (réel)	Obligation crédit	Actions	immobilier
Monétaire	100,00 %	-7,66 %	-8,55 %	-9,57 %	0,84 %	0,78 %	-7,66 %	-0,21 %	1,82 %
ZC 5 ans (nominal)	-7,66 %	100,00 %	98,56 %	97,31 %	56,54 %	55,46 %	100,00 %	-38,98 %	1,01 %
ZC 10 ans (nominal)	-8,55 %	98,56 %	100,00 %	99,73 %	46,92 %	46,79 %	98,56 %	-35,82 %	1,31 %
ZC 15 ans (nominal)	-9,57 %	97,31 %	99,73 %	100,00 %	42,77 %	42,91 %	97,31 %	-32,13 %	0,97 %
ZC 8 ans (réel)	0,84 %	56,54 %	46,92 %	42,77 %	100,00 %	99,68 %	56,54 %	-44,60_%	5,21 %
ZC 15 ans (réel)	0,78 %	55,46 %	46,79 %	42,91 %	99,68 %	100,00 %	55,46 %	-44,22 %	5,32 %
Obligation crédit	-7,66 %	100,00 %	98,56 %	97,31 %	56,54 %	55,46 %	100,00 %	-38,98 %	1,01 %
Actions	-0,21 %	-38,98 %	-35,82 %	-32,13 %	-44,60 %	-44,22 %	-38,98 %	100,00 %	-10,57 %
Immobilier	1,82 %	1,01 %	1,31 %	0,97 %	5,21 %	5,32 %	1,01 %	-10,57 %	100,00 %

Tab. 25 : *Corrélations obtenues lors des projections entre les différents rendements dans un régime à forte volatilité*

Les graphiques suivants présentent l'évolution des rendements des actions entre les deux régimes sur une trajectoire simulée (choisie à titre illustratif). Le nombre de fois où ces rendements se trouvent sous le régime 2 (régime de « crise ») est nettement plus faible que celui sous le régime 1 (régime normal). Cela reflète une probabilité plus faible de se retrouver dans le régime de crise, comme supposé dans les hypothèses de départ dans le modèle des actions.

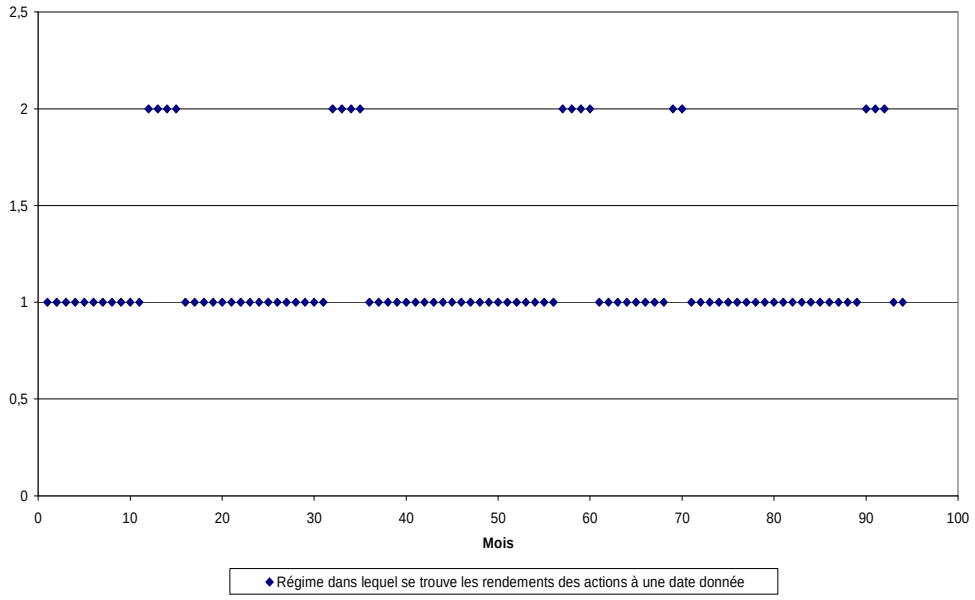


Fig. 30 : *Evolution du marché des actions entre le régime 1 (actions à faible volatilité) et le régime 2 (actions à forte volatilité) sur une trajectoire simulée*

Comparaison de l'approche du coefficient d'abattement avec l'approche « mélange » pour le crédit

Nous avons montré que les approximations retenues pour la méthode « mélange » sont délicates à justifier. Au-delà des descriptions fonctionnelles de cette approche, il s'agit ici de présenter quelques résultats obtenus en l'utilisant.

L'approche du coefficient d'abattement consiste à calculer le prix de l'obligation crédit en multipliant le prix des obligations nominales par un coefficient d'abattement dépendant de la maturité.

Aussi, dans un premier temps le tableau ci-après compare les rendements et les volatilités des obligations du crédit correspondant aux deux approches : l'approche du coefficient d'abattement et l'approche « mélange ». Notons que la discrétisation des formules, utilisée lors du calcul, est exacte dans les deux approches. Le calibrage retenu pour cette comparaison est le même que celui présenté dans la sous-section II-6 du chapitre 2 à l'exception de la matrice de corrélation des browniens qui intègre en plus une ligne et une colonne correspondant aux corrélations avec les *spreads* de crédit.

Obligation crédit 5 ans	**Coefficient d'abattement**	**approche « mélange »**
Rendements	4,27 %	3,98 %
Volatilités	4,42 %	4,27 %

Tab. 26 : *Discrétisation exacte pour Coefficient d'abattement et l'approche « mélange »*

Nous observons que les rendements et les volatilités sont du même ordre de grandeur. Cela signifie que les approches, bien que différentes, conduisent à des résultats proches. Ainsi, l'approche du coefficient d'abattement donne des résultats comparables à ceux de l'approche « mélange » en termes de rendements moyens et de volatilités.

Effet de la discrétisation

À titre illustratif, le tableau suivant présente les rendements et les volatilités des deux approches mais en retenant la discrétisation d'Euler pour toutes les variables modélisées dans le cas de l'approche « mélange » (nous rappelons que toutes les variables sont implémentées en tenant compte d'une discrétisation exacte dans le cas de l'autre approche des coefficients d'abattement). En termes de résultat, nous constatons que les rendements moyens dans les deux approches sont proches. Cependant les volatilités sont significativement différentes. Donc le choix de la discrétisation a un impact significatif sur le niveau de la volatilité des obligations crédit lorsqu'elles sont modélisées avec ces deux approches différentes.

Obligation crédit 5 ans	**Discrétisation exacte**	**Discrétisation d'Euler**
Rendements	4,27 %	3,96 %
Volatilités	4,42 %	7,43 %

Tab. 27 : *Discrétisation exacte pour Coefficient d'abattement et d'Euler pour l'approche « mélange »*

Conclusion du chapitre 2

L'objet du chapitre 2 de cette partie était de présenter un modèle stochastique de génération des scénarios économiques dont la vocation finale est plus proche du domaine de la gestion des risques que de celui du *pricing*. Nous avons également essayé d'étudier différents aspects liés au processus de calibrage du GSE construit et de mettre en avant les éléments pertinents relatifs à ce processus.

Nous rappelons qu'il existe, dans la littérature, d'autres modèles candidats (cf. Ahlgrim et al. [2008]). Nous ne prétendons pas que le modèle présenté est le « meilleur » modèle (d'ailleurs aucun modèle ne l'est) mais nous avons la garantie que celui-ci respecte certaines conditions minimales telles que la parcimonie et la plausibilité économique dans le sens où le GSE utilisé est capable d'imiter certains phénomènes financiers importants observés sur les marchés financiers. Ces phénomènes attirent souvent l'attention des analystes qui s'intéressent aux comportements joints de l'inflation, des taux d'intérêts et des marchés des actions. Les états de sortie (ou l'*output*) du modèle proposé ont été ainsi illustrés en partant initialement des paramètres obtenus via l'approche de calibrage décrite.

Conclusion de la partie I

Le glissement sémantique du « modèle d'actifs » vers le « générateur de scénarios économiques » matérialise la profonde évolution de ce sujet depuis une dizaine d'années maintenant.

Longtemps utilisée dans le cadre d'études spécifiques et ponctuelles, la construction d'une représentation des actifs dans lequel l'assureur peut investir ses avoirs et du contexte économique dans lequel ils évoluent est devenue un élément essentiel de la description des risques que celui-ci doit gérer.

Que ce soit dans le cadre prudentiel, pour la détermination des provisions et du capital de solvabilité, pour sa communication vers les tiers (MCEV et états comptables) ou pour ses besoins de pilotage techniques (choix d'allocations stratégique et tactique, tests de rentabilité, etc.), l'organisme assureur doit disposer d'un cadre rigoureux et cohérent prenant en compte l'ensemble des actifs de son bilan et les risques associés.

Si l'identification de ces risques peut être considérée comme relativement aboutie, la crise financière a mis en évidence certaines faiblesses dans leur modélisation. Deux éléments sont ainsi mis en évidence et vont sans doute donner lieu à de nombreux développements dans les prochaines années (cf. Planchet et al. [2009]) :

- la structure de dépendance entre les actifs ;
- le risque de liquidité, intimement associé à la gestion efficace des couvertures financières.

La quasi-totalité des modèles actuels s'appuient sur des structures de dépendance dans lesquels la corrélation tient une place centrale ; de nombreux travaux, dont nous pourrons trouver une synthèse dans Kharoubi-Rakotomalala [2008], mettent en évidence le caractère dynamique de l'intensité de la dépendance. En pratique, l'intensité de la dépendance augmente dans les situations défavorables, ce qui limite l'efficacité des mesures de

diversifications calibrées avec des structures ne prenant pas en compte cet effet. L'introduction de structures de dépendance non linéaires intégrant de la dépendance de queue apparaît ainsi comme un élément incontournable de l'évolution des générateurs de scénarios économiques.

Le risque de liquidité est également apparu comme un élément majeur de la crise des *subprimes* et, plus généralement, de la crise financière qu'elle a engendrée. Au moment où la généralisation des approches *market consistent* (basées sur les valeurs de marché) impose d'évaluer les options et garanties financières des portefeuilles dans la logique de détermination du coût de leur couverture, encore faut-il pour que le montant obtenu ait du sens que la couverture puisse être réajustée régulièrement, ce qui n'est possible qu'avec des actifs liquides. Cela impose par conséquent si ce n'est une adaptation de l'approche dans le cas d'actifs peu liquides, à tout le moins la prise en compte d'une prime de liquidité pour refléter dans le montant affiché ce risque d'impossibilité de gérer idéalement la couverture.

Au-delà de ces deux éléments structurants, l'efficacité opérationnelle des modèles mis en œuvre dépend, nous l'avons vu, dans une très large mesure de la pertinence des paramètres retenus pour les alimenter. La détermination de ses paramètres est complexe et fait appel à la fois à des considérations d'ordre statistique (exploitation des historiques), économique (cohérence des valeurs de long terme prédites par le modèle avec les relations économiques fondamentales), financières enfin (cohérence avec les prix observés sur le marché). La prise en compte rationnelle de ces différentes composantes nécessite une réflexion spécifique et fait partie intégrante des choix structurants en termes de gestion des risques que peut effectuer l'organisme assureur.

Enfin, la définition d'indicateurs de performance du modèle contribue à une meilleure alimentation du processus de prise de décision. L'exploitation de ces indicateurs dans le cas de modèles plus compliqués et plus concrets, tel que les modèles de gestion actif-passif, constitue notre prochain champ d'investigation.

PARTIE II :

L'ALLOCATION STRATÉGIQUE D'ACTIFS DANS LE CADRE DE LA GESTION ACTIF-PASSIF

Introduction

Cette seconde partie est consacrée à l'élaboration de l'allocation d'actif elle-même. Rappelons que concernant les régimes de retraite par répartition, trois situations peuvent être recensées : régime provisionné (cas où le régime est tenu à la couverture totale des engagements souscrits par ses cotisants actuels et futurs), régime partiellement provisionné (couverture d'une partie seulement des engagements souscrits par ses cotisants actuels et futurs) et enfin régime non provisionné.

Le fonctionnement financier des régimes de retraite en France (objet de la plupart des applications illustratives dans ce travail) peut être schématisé comme suit : les flux de cotisations permettent de régler les flux de prestations. Le surplus, le cas échéant, permet d'alimenter une réserve destinée à régler une partie des prestations futures. Cette même réserve peut se voir prélever, le cas échéant, le solde technique débiteur (cotisations insuffisantes pour régler les prestations). Entre temps, la réserve est placée sur le marché financier et répartie entre différentes classes d'actifs.

La gestion actif-passif d'un régime par répartition partiellement provisionné peut être basée sur l'optimisation de la valeur de la réserve, compte tenu des contraintes liées au passif qu'il doit respecter. C'est dans ce cadre que le choix d'une « bonne » allocation stratégique joue un rôle essentiel dans le pilotage actif-passif d'un régime de retraite, étant donné que les réserves contribuent à part entière à la solidité du régime.

Cependant une difficulté consiste à définir la « meilleure » stratégie de placement de ces réserves sur les marchés financiers, notamment dans un contexte de fortes incertitudes économiques, comme c'est actuellement le cas du fait de la récente crise financière.

Il est à rappeler que le modèle d'allocation d'actifs le plus répandue est celui de Markowitz [1952]. L'objectif de ce modèle est de résoudre le problème suivant : pour une rentabilité espérée fixée par l'investisseur, trouver les proportions à investir dans chaque titre, conduisant à un portefeuille de risque minimum et sous un ensemble de contraintes. L'ensemble des portefeuilles vérifiant le programme d'optimisation de Markowitz pour différents niveaux de rentabilité espérée, décrit une hyperbole dans le repère (risque mesuré par l'écart type de la rentabilité / rentabilité espérée).

Le modèle de Markowitz n'est pas un modèle de gestion au jour le jour mais un modèle de structure à long terme qui a conduit à de nombreux développements sur la théorie de la gestion de portefeuille. La limite de ce modèle réside dans le fait d'utiliser la variance pour mesurer le risque du portefeuille. Quoique simple et parfois efficace, la variance est contre intuitive : Les investisseurs n'ont aucune aversion face à des rendements supérieurs à ceux prévus, ils s'intéressent seulement à ce qui est en dessous de leurs prévisions où de leurs objectifs personnels. Or, la variance est une mesure de dispersion qui ne fait aucune différence entre ce qui est en dessous et au-dessus de l'espérance de rendement. En plus il ne tient pas compte de l'existence d'un passif.

Dans la littérature, les modèles d'ALM sont généralement classés en trois groupes présentés chronologiquement comme suit.

Le premier groupe contient les modèles d'adossement (ou *matching*) et d'immunisation par la duration (cf. Macaulay [1938], Redington [1952]). Ces modèles se basent sur le fait que les

investissements sont essentiellement effectués dans des obligations. Ceci nous permet d'obtenir, soit un adossement des flux de trésorerie des actifs financiers à ceux du passif (*matching)*, soit un adossement de la duration de l'actif à celle du passif (immunisation par la duration).

Ces techniques étaient utilisées jusqu'au milieu des années 80 et avaient comme inconvénients principaux la considération du risque de taux comme seule source de risque pour le fonds, ainsi que la nécessité d'un rebalancement périodique du portefeuille en ré-estimant à chaque fois la duration du passif, qui change continûment du fait du changement des taux d'intérêts et de l'écoulement du temps. Pierre [2009] présente une approche de la couverture de passif qui vise à résoudre les problèmes mentionnés ci-dessus. L'actif sera constitué dans ce cas d'un portefeuille de couverture de type taux (obligations ou dérivés) couplé à un portefeuille de rendement. Selon Pierre [2009], cette architecture permet une flexibilité suffisante pour permettre à l'actif de s'adapter aux mise à jour de la valeur des engagements lors de la revue des hypothèses actuarielles ayant permis leur évaluation.

Le deuxième groupe contient les modèles basés sur la simulation de scénarios déterministes et sur la notion de surplus (Kim et Santomero [1988], Sharpe et Tint [1990], Leibowitz et al. [1992]). Les modèles de surplus ont pour objet la minimisation du risque de perte du surplus (mesuré par la variance de la rentabilité du surplus) sous contraintes de rentabilité et de poids des actifs. Ils sont des modèles mono-périodiques, ce qui limite leur utilité en pratique pour des problèmes d'allocation sur le long terme.

Le troisième groupe de modèles utilise les techniques de simulation stochastique (*Monte Carlo*) pour modéliser l'évolution des différents éléments, que ce soit au niveau des actifs financiers et des engagements, ou au niveau des variables de marché et des variables démographiques (cf. Frauendorfer [2007], Munk et al. [2004], Waring [2004], Martellini [2006]). Ainsi les lois de probabilité associées aux résultats du fonds de retraite sur le long terme peuvent être estimées. A ce niveau, nous nous proposons de distinguer deux sous-groupes de modèles d'ALM basés sur les techniques stochastiques. L'élément clé de distinction sera si oui ou non les poids des différents actifs reviennent périodiquement à ceux de l'allocation stratégique définie initialement (si oui, les modèles seront appelés modèles à poids constants ou stratégie *Fixed-Mix*).

- Pour le premier sous-groupe de modèles à poids constants et malgré les avancées réalisées avec ces techniques (surtout au niveau de l'implémentation informatique), l'aspect dynamique de l'allocation stratégique reste encore marginalisé. En fait, ces modèles permettent de comparer des allocations constantes dans le temps (statiques) indépendamment des opportunités liées aux évolutions inter-temporelles des marchés (cf. Merton [1990], Kouwenberg [2001], Infanger [2002], Dempster et al. [2003]).

- Le deuxième sous-groupe de modèles, et le plus récent, est principalement inspiré de la théorie du choix de la consommation et de portefeuille développée par Merton [1971]. Il s'agit des modèles d'allocation dynamique ou inter temporels. Par exemple, à partir de la définition de la fonction objectif pour l'investisseur, ces modèles permettent la détermination d'une trajectoire des poids des différents actifs jusqu'à la date d'échéance (l'ajustement des poids est fonction des évolutions projetées du marché et de la règle de gestion prédéfinie). L'allocation stratégique retenue sera l'allocation optimale d'actifs

à la date initiale t_0. Le cadre d'utilisation de ces modèles récents se heurte au problème d'implémentation vu la complexité des outils mathématiques employés (cf. Hainaut et al. [2005], Rudolf et al. [2004] et Yen et al. [2003]).

Concernant les modèles d'allocation d'actifs, notre étude est axée sur la comparaison des modèles disponibles selon les hypothèses sous-jacentes de « rebalancement » des portefeuilles : nous faisons la distinction entre une gestion « statique » (pour laquelle les poids reviennent périodiquement à ceux de l'allocation stratégique de long-terme - cas par exemple des modèles à poids constants *Fixed-Mix*) et une gestion dite « dynamique », pour laquelle les poids peuvent s'écarter définitivement de l'allocation stratégique initiale selon des règles de gestion prédéfinies.

Nous verrons que l'approche dynamique présente l'avantage théorique de la robustesse face aux changements de régime des marchés. L'autorisation du changement des poids des différentes classes d'actifs, sur la base d'une règle de gestion bien définie, constitue *a priori* un élément intéressant. Cela en effet permet l'ajustement des expositions aux différentes classes d'actifs suite à l'évolution des conditions de marché.

Cette réflexion sur la construction de modèles d'allocation d'actifs applicables dans une optique prévisionnelle à long-terme nous conduira à étudier une approche innovante fondée sur les techniques de « programmation stochastique » (cf. Birge et Louveaux [1997]). Il s'agit d'une version adaptée d'une technique déjà utilisée dans le domaine de l'ingénierie pour la planification de la production (cf. Dantzig et al. [1990], Escudero et al. [1993]). L'objectif sera la mise en évidence et l'étude des caractéristiques de cette approche.

Comme déjà mentionné, cette thèse accorde également une attention toute particulière aux techniques numériques de recherche de l'optimum, qui demeurent des questions essentielles pour la mise en place d'un modèle d'allocation. Le point de départ sera notre constat d'un temps de calcul significatif dû simultanément à un nombre élevé de scénarios économiques générés et à un nombre d'allocations d'actifs testées également élevé.

Dans ce cadre, nous présentons un algorithme d'optimisation globale d'une fonction non convexe et bruitée. L'algorithme est construit après une étude de critères de compromis entre, d'une part, l'exploration de la fonction objectif en de nouveaux points (correspondant à des tests sur de nouvelles allocations d'actifs) et d'autre part l'amélioration de la connaissance de celle-ci, par l'augmentation du nombre de tirages en des points déjà explorés (correspondant à la génération de scénarios économiques supplémentaires pour les allocations d'actifs déjà testées). Une application numérique illustre la conformité du comportement de cet algorithme à celui prévu théoriquement.

Le plan de cette deuxième partie sera comme suit :

- Premier chapitre : Il est question dans ce chapitre d'analyser les approches classiques (ou déterministes) de gestion actif-passif à savoir les modèles basés sur la notion de « duration » et de « surplus ». Nous détaillons différentes approches tout en mettant en évidence la différence entre elles. L'objectif de ce chapitre est de revoir l'état de l'art en matière de modèles d'ALM déterministe. Après avoir mis en évidence les limites de ces modèles, nous passons à l'étude de modèles plus élaborés et plus sophistiqués (objet des chapitres ultérieurs).

- Deuxième chapitre : Nous considérons ici une approche récente d'allocation stratégique d'actifs basée sur la stratégie dite « à poids constants » ou *Fixed-Mix* (cf. Merton [1990], Kouwenberg [2001], Infanger [2002], Dempster et al. [2003]).

Nous proposons une modélisation du régime-type de retraite et étudions certains critères d'allocation stratégique d'actifs en fonction du type du régime (provisionné, partiellement provisionné, etc.) sont étudiés. Nous passons ensuite à l'illustration de la stratégie *Fixed-Mix* avec une application sur les réserves d'un régime de retraite partiellement provisionné. Les résultats obtenus sont discutés et différents tests de sensibilité sont mises en place : ces tests sont liés principalement à l'impact des hypothèses de rendement ou de corrélation retenues.

Les difficultés rencontrées lors de la mise en place de la stratégie *Fixed-Mix,* dues essentiellement à la multiplicité du nombre de classes d'actifs considérées et au nombre de scénarios économiques simulés, nous ont mené à nous pencher sur les aspects d'optimisation numérique. Le point de départ est notre constat d'un temps de calcul significatif dû simultanément à un nombre élevé de scénarios économiques générés et à un nombre d'allocations d'actifs testées également élevé.

Dans ce cadre, une présentation détaillée des travaux menés lors de la rédaction de l'article de Rullière et al. [2010] est effectuée. Un algorithme d'optimisation globale d'une fonction non convexe et bruitée est présenté. L'algorithme est construit après une étude de critères de compromis entre, d'une part, l'exploration de la fonction objectif en de nouveaux points (correspondant à des tests sur de nouvelles allocations d'actifs) et d'autre part l'amélioration de la connaissance de celle-ci, par l'augmentation du nombre de tirages en des points déjà explorés (correspondant à la génération de scénarios économiques supplémentaires pour les allocations d'actifs déjà testées). Une application numérique illustre la conformité du comportement de cet algorithme à celui prévu théoriquement et compare les résultats obtenus avec l'algorithme de Kiefer-Wolfowitz- Blum (cf. Blum [1954], Kiefer et Wolfowitz [1952]).

- Troisième chapitre : Les modèles classiques d'ALM dynamiques sont explorés, notamment les techniques d'assurance de portefeuille (cf. Perold et Sharpe [1988]) et les techniques de programmation dynamique (cf. Cox et Huang [1989], Merton [1971]). A ce niveau, les techniques d'assurance de portefeuille basées sur la notion de CPPI ou *Constant Proportion Portfolio Insurance* (cf. Perold et Sharpe [1988]) sont mises en place et certains résultats relatifs à ce modèle sont étudiés. De même, les principes des techniques de programmation dynamique et leurs limites sont également mises en évidence.

Nous nous penchons par la suite sur une approche innovante fondée sur les techniques de « programmation stochastique » (cf. Birge et Louveaux [1997]). Il s'agit d'une version adaptée d'une technique déjà utilisée dans le domaine de l'ingénierie pour la planification de la production (cf. Dantzig et al. [1990], Escudero et al. [1993]). Dans ce cadre, nous mettons en place un modèle d'ALM dynamique basé sur les techniques de programmation stochastique.

Nous proposons, au cours d'une illustration numérique, une nouvelle méthodologie de génération de scénarios économiques que nous appelons méthodologie « des quantiles de référence ». Cette dernière permet de partir d'une structure linéaire de génération de scénarios (telle que décrite dans le chapitre 2 de la partie I) pour réduire la dimension du problème rencontré avec la stratégie *Fixed-Mix* tout en tenant compte de la corrélation entre les distributions des différentes variables projetées.

A travers la même application numérique, nous comparons certains résultats relatifs aux deux approches d'allocation stratégique d'actifs : celle basée sur la stratégie *Fixed-Mix* et celle basée sur les techniques de programmation stochastique. Nous testons également la sensibilité de cette deuxième approche par rapport au changement de certains de ses paramètres, toutes choses étant égales par ailleurs.

Chapitre 1 : Les modèles d'ALM classiques (déterministes)

Il est question dans ce qui suit d'analyser les approches classiques (ou déterministes) de gestion actif-passif à savoir les modèles basés sur la notion de « duration » et de « surplus ». Nous détaillons différentes approches tout en mettant en évidence la différence entre elles. L'objectif de ce chapitre est de revoir l'état de l'art en matière de modèles d'ALM déterministe.

I- Immunisation du portefeuille

Redington [1952] définit l'immunisation du portefeuille comme « l'investissement de l'actif d'une telle manière que le portefeuille soit protégé contre un changement des taux d'intérêt ». Autrement, c'est une stratégie d'investissement qui, dans le cas de l'assurance vie ou de celui des régimes de retraite, produit des flux exactement adossés en maturité et en valeur à ceux que doit payer l'entreprise. Cette technique suppose que le risque principal des portefeuilles financiers est celui du changement des taux d'intérêt.

Nous pouvons énumérer deux approches classiques appartenant à la catégorie de l'immunisation :

- l'adossement des flux de trésorerie (*cash flows*) à ceux du passif.
- l'adossement des durations de l'actif et du passif.

Nous rappelons cependant quelques notions qui seront reprises ultérieurement dans cette étude (cf. Le Vallois et al. [2003]).

Taux de rendement actuariel

En suivant à une démarche inverse à celle du calcul de la valeur actuelle d'un actif (ou de toutes séquences de flux fixes), il est possible de calculer, à partir du prix de marché, un taux d'actualisation des flux de trésorerie correspondants dit taux de rendement actuariel r_a. Ce dernier vérifie :

$$Prix\ de\ marché = Valeur\ actuelle = \sum_{i=1}^{n} \frac{F_{t_i}}{(1+r_a)^{t_i}}$$

Avec F_{t_i} est le flux de l'actif financier (coupon, dividende, amortissement…) à l'époque t_i.

Courbe des taux zéro-coupon

Le calcul du taux de rendement actuariel des actifs financiers, et des obligations en particulier, présente l'inconvénient de ne pas donner la même valeur de taux pour les différents actifs. Dans le cas des obligations, cela est dû principalement à deux facteurs.

- Le premier facteur est que les investisseurs demandent aux émetteurs privés des primes de risque appelés *spread* de signature qui compense le risque de défaillance propre à chaque émetteur (ce *spread* diffère donc d'une obligation à l'autre).

- Le deuxième facteur est la dépendance des taux de rendement actuariels des échéances des obligations correspondantes. En fait même si nous limitons l'analyse aux emprunts d'Etat, tous identique en terme de risque, nous

constatons encore des différences entre les taux de rendement actuariels de différents titres. Il faudrait utiliser un taux d'actualisation différent pour chaque échéance. Le seul cas ou le taux de rendement actuariel observé correspond sans ambiguïté à une seule maturité est le cas où il n'y a qu'un seul flux de trésorerie. Ce cas existe puisqu'il correspond aux titres dits obligations zéro-coupon pour lesquels les intérêts sont versés en une seule fois au moment de l'échéance finale et unique.

L'observation des prix des zéro-coupon permettrait donc de bâtir une courbe des taux zéro-coupon. Cette courbe existe, mais en pratique elle est obtenue par des moyens différents compte tenu de la faible liquidité des zéro-coupon existants sur le marché. La courbe zéro-coupon est également appelée structure par terme des taux, ou encore courbe des taux spot.

Valeur actuelle nette des actifs obligataires

En pratique il n'est pas nécessaire de calculer la valeur actuelle des actifs obligataires, puisqu'il est plus simple d'observer leur valeur de marché. Cependant, dans le cas des obligations à taux fixe, la séquence des flux de trésorerie futurs associés à un titre est parfaitement connue.

Nous considérons même généralement que dans le cas des emprunts d'Etat, cette séquence est certaine et ne présente aucun risque de défaillance de l'émetteur. A l'aide de la structure par terme des taux (notée r_{t_i}) il est possible de reconstituer le prix d'équilibre de l'obligation sans risque : $Prix\ de\ marché \approx Prix\ estimé = \sum_{i=1}^{n} \frac{F_{t_i}}{\left(1+r_{t_i}\right)^{t_i}}$

Valeur actuelle des passifs

Par analogie avec la valorisation des actifs, la valorisation des flux du passif se fait en se basant sur la courbe des taux zéro-coupon. Cette dernière, permet de tenir compte de la maturité de chacun des flux pour lui attribuer un taux d'actualisation précis. De même, il faudrait ajouter à cette courbe des taux sans risque une prime de risque (*spread*). Elle permettra de retomber sur une « valeur de marché » des passifs. Le problème c'est que le niveau réel de cette prime est inobservable sur le marché dans la mesure où les engagements du passif ne s'échangent pas régulièrement sur des marchés liquides et organisés. En pratique, le calcul de la valeur actuelle du passif peut être réalisé avec une prime de risque arbitrairement choisie, éventuellement nulle.

I-1 Adossement des flux de trésorerie

Il s'agit de la procédure d'immunisation la plus simple et la plus ancienne. Elle consiste à investir la richesse initiale dans un portefeuille de titres (le plus souvent des zéro-coupons) qui produisent exactement, et aux échéances prévues, les flux du passif.

Autrement, lorsque sur chaque période les flux nets obtenus (total des entrées de flux –total des sorties de flux) sont toujours positifs ou nuls, l'actif est dit adossé au passif. Il est dit exactement ou parfaitement adossé si les flux nets sont nuls. Les anglo-saxons parlent dans ce cas de méthode de *cash flow matching*. Une fois que tous les passifs sont couverts par cette méthode, les actifs excédentaires sont alors considérés comme libres et représentatifs de la

situation nette réelle de la société (le *shareholder surplus*). Le même raisonnement peut être adopté dans le cas des régimes de retraite par répartition.

En pratique, l'adossement par les flux de trésorerie est toujours utilisé par les sociétés avec la différence que les passifs peuvent être regroupés sur d'autres critères que le seul « rendement ». En fait, les engagements au passif des assureurs par exemple sont généralement répartis par famille de contrats (en fonction des prix de revient, des taux de rendement actuariels, de la durée des titres…) appelés contons, puis les provisions correspondantes et les flux de trésorerie associés sont calculés. Cela permet outre l'adossement des flux de trésorerie, la bonne adéquation des actifs en termes de taux de rendement financier et comptable comparé au taux garanti moyen du passif correspondant.

L'adossement des flux de trésorerie est valable à très court terme, mais malheureusement, elle ne trouve pas d'applications pratiques à long terme. En effet, les flux de l'actif et ceux du passif sont très influencés par des facteurs externes, notamment les taux d'intérêt. Ils sont donc eux-mêmes sujet à des variations dans des sens différents, il importe donc de réitérer souvent cette immunisation.

I-2 Adossement par la duration

Cette technique consiste à apparier les sensibilités de l'actif et du passif vis-a- vis de la variation des taux d'intérêt. En d'autres termes, elle consiste à définir une stratégie d'investissement qui fait que la valeur de marché des actifs suit tout mouvement de la valeur actuelle des engagements. L'immunisation par la duration définit un portefeuille dont la valeur, au premier ordre, évolue comme la valeur actuelle des engagements. La règle de décision dans cette technique est basée sur l'indice de « sensibilité », défini par Macauley [1938]. Il est obtenu à partir de la formule de développement limite de Taylor du prix en fonction du taux d'intérêt.

Pour un titre, dont les caractéristiques contractuelles sont déterminées indépendamment du mouvement des taux, la sensibilité est exprimée comme la variation relative de cette valeur induite par une variation infinitésimale de taux d'intérêt.

- La sensibilité de l'actif :

Nous reprenons l'expression simplifiée de la valeur actuelle d'une obligation à taux fixe en fonction du taux d'actualisation actuariel r_a :

$$VA(r_a) = \sum_{i=1}^{n} \frac{F_{t_i}}{(1+r_a)^{t_i}}$$

La variation de la valeur actuelle pour une petite variation du taux d'actualisation est donnée par le calcul de la dérivée première :

$$\frac{dVA(r_a)}{dr} = VA'(r_a) = \sum_{i=1}^{n} -t_i \frac{F_{t_i}}{(1+r_a)^{t_i+1}}$$

Cette dérivation n'a de sens que si les flux F_{t_i} sont fixes par rapport à r .

La sensibilité est ainsi donnée par l'expression suivante :

$$sensibilité = \frac{dVA(r_a)}{VA \times dr} = -\frac{1}{VA} \times \sum_{i=1}^{n} t_i \frac{F_{t_i}}{(1+r_a)^{t_i+1}}$$

Lorsque les flux de trésorerie sont tous positifs, la sensibilité de la valeur actuelle aux variations du taux d'actualisation est nécessairement négative. Cela est effectivement le cas pour les obligations dont la valeur de marché croît quand les taux baissent et réciproquement. La sensibilité est appelée aussi duration modifiée.

- La duration de l'actif

La duration telle que définie par Macaulay [1938] est donnée par l'expression suivante :

$$duration = \frac{\sum_{i=1}^{n} t_i \frac{F_{t_i}}{(1+r_a)^{t_i}}}{\sum_{i=1}^{n} \frac{F_{t_i}}{(1+r_a)^{t_i}}} = \frac{\sum_{i=1}^{n} t_i \frac{F_{t_i}}{(1+r_a)^{t_i}}}{VA(r_a)}$$

Avec F_{t_i} une série de flux fixes.

La duration peut s'interpréter comme la durée de vie moyenne de l'obligation. En fait, chaque durée t_i étant pondérée par la valeur actuelle du flux correspondant $\frac{F_{t_i}}{(1+r_a)^{t_i}}$. Elle peut s'interpréter aussi comme l'élasticité du prix de l'obligation par rapport aux variations des taux. C'est la définition de Hicks [1946] qui relie la sensibilité à la duration :

$$duration = -\frac{\frac{dVA(r_a)}{VA(r_a)}}{\frac{dr_a}{1+r_a}} = -(1+r_a) \times \frac{\frac{dVA(r_a)}{VA(r_a)}}{dr_a} = -(1+r_a) \times sensibilité$$

La duration s'interprète aussi comme la date à la quelle les deux effets de la variation des taux, effet sur la valeur et effet sur le revenu, se compensent.

Elle est compatible avec une structure par terme des taux à condition de se limiter aux variations parallèles des courbes de taux. Autrement, la variation de taux dr_t doit s'appliquer d'une façon uniforme à la courbe des taux (chaque r_{t_i} devient $r_{t_i} + dr_t$) ce qui correspond bien à une translation uniforme de la courbe des taux. La duration s'exprime dans ce cas comme suit :

$$duration = \frac{\sum_{i=1}^{n} t_i \frac{F_{t_i}}{(1+r_{t_i})^{t_i}}}{\sum_{i=1}^{n} \frac{F_{t_i}}{(1+r_{t_i})^{t_i}}} = \frac{\sum_{i=1}^{n} t_i \frac{F_{t_i}}{(1+r_{t_i})^{t_i}}}{VA(r_{t_i})}$$

Si nous étendons la notion de la duration à des flux variables (obligations à taux variables par exemple), nous pourrons dans certains cas obtenir la relation ci dessus entre la duration et la sensibilité (tant que les dérivées $F'_{t_i}(r_a)$ sont connues), mais la différence c'est que la duration n'est plus assimilable à une durée moyenne.

La duration peut être négative si certains des flux sont négatifs. La sensibilité dans ce cas est positive : la valeur actualisée de la séquence examinée augmente quand les taux augmentent et inversement. La duration d'une obligation zéro coupon est égale à la durée de cette obligation.

Le calcul de la duration d'un portefeuille s'effectue par trois méthodes :

- La première consiste à déterminer la moyenne des durations pondérées par la valeur de marché des titres correspondant (coupons courus inclus). L'inconvénient de cette méthode c'est qu'elle ne peut avoir de sens que si les taux actuariels de tous les titres étaient identiques, ce qui bien entendu n'est jamais le cas.

- La deuxième méthode consiste à cumuler tous les flux de trésorerie du portefeuille et à calculer un taux de rendement actuariel unique en fonction de la valeur de marché totale, coupon courus inclus. Il est ensuite possible de calculer une duration globale puisque la formule de calcul est valable à toute séquence de flux fixes.

- La troisième méthode est la plus rigoureuse. Elle consiste à calculer la valeur actuelle et la duration de chaque titre avec une courbe des taux commune, puis à calculer la moyenne pondérée par les valeurs actuelles et non pas par les valeurs de marché. C'est théoriquement la meilleure solution, car chaque flux de trésorerie est actualisé avec le taux qui correspond à sa maturité. En pratique, la différence obtenue entre les différentes méthodes de calcul de la duration d'un portefeuille n'est pas significative.

Par ailleurs, nous pouvons étendre la notion de duration à d'autres actifs non obligataires malgré qu'ils sont plus ou moins sensibles à la variation des taux notamment les actions. Pour ces derniers le calcul de la duration peut être effectué dans le cadre du modèle Gordon-Shapiro [1956] (qui permet d'analyser le prix d'une action comme étant l'actualisation de la série des dividendes projetés) ou tout simplement dans le cadre de la régression des prix de marché par rapport aux variations de taux. Les actifs et les passifs conditionnels ne doivent donc pas faire l'objet de calculs élémentaires d'adossement en duration.

Pour accomplir une immunisation par la duration, l'investisseur doit acquérir des titres dont la duration moyenne est égale à la duration des flux du passif.

La notion de sensibilité est un indicateur de l'exposition au risque de taux par la Valeur Actuelle Nette (VAN) de la société.

En fait, à partir de la relation suivante : $VAnette = VAactif - VApassif$ nous pouvons exprimer l'exposition de la VAN au risque de taux comme suit :

$$\frac{dVAnette}{dr} = \frac{dVAactif}{dr} - \frac{dVApassif}{dr}$$

Si nous considérons le cas où la VAN initiale est égale à zéro, et où la sensibilité de l'actif et du passif sont identiques, alors la variation de taux est sans influence sur la VAN :

$$\frac{dVAnette}{dr} = 0$$

Nous parlons dans ce cas de l'immunisation du risque de taux. En effet, la VAN devient insensible, sinon à toute variation de taux, du moins à une petite variation parallèle de la courbe des taux. Cet effet est obtenu en alignant les sensibilités de l'actif et du passif.

Nous notons également que :

- Plus la duration est élevée plus le risque est grand et plus son prix est sensible aux variations du taux de marché.

- Plus le coupon est faible, plus la sensibilité est élevée.

- Nous couvrons systématiquement le passif par une obligation de durée plus longue (car la duration est inférieure à la durée sauf pour le cas des zéro-coupon).

- En cas de baisse des taux, le gain réalisé sur la valeur de l'obligation est supérieur à la perte encourue sur le passif.

- L'immunisation en duration d'un portefeuille n'est parfaite que si elle est réalisée avec des instruments zéro-coupon. Elle ne tient pas compte de la forme de la courbe des taux, ni de ses déformations, ni de sa dynamique.

- Limites de l'adossement par la duration

Ils existent deux limites principales des indicateurs ci-dessus :

- Le domaine d'utilisation de ces concepts est limité aux variations parallèles de la courbe des taux.
- Les calculs de la duration et de la convexité ne gardent leurs significativité que pour des flux fixes, indépendants des taux de marché.

La première limite n'est pas sans conséquence mais elle n'invalide pas totalement l'analyse du risque de taux. En effet, les variations parallèles de la courbe des taux représentent la principale source de risque pour la plupart des portefeuilles obligataires (résultats obtenus par des analyses statistiques de la variance des taux en composante principale).

En revanche, la deuxième limite (flux fixes) est très contraignante pour l'analyse des passifs comprenant des options complexes exercées par les clients (tel que l'option de rachat de l'épargne). L'exercice de certains de ces options dépend fortement de l'évolution des taux sur le marché. Pour le cas d'option de rachat, une augmentation des niveaux de taux a pour conséquence évidente la hausse des niveaux de rachats afin de profiter de placement plus rentable sur le marché. En pratique, nous faisons souvent l'hypothèse que pour de petites variations des taux de marché, les flux de trésorerie liés au passif restent fixes. Les calculs de

sensibilité et de duration ont alors une validité « locale » dont l'intérêt est cependant limité pour l'adéquation actif-passif.

Par ailleurs, l'immunisation par la duration requiert un rebalancement périodique du portefeuille en réestimant à chaque fois la duration du passif, celle-ci change continûment du fait du changement des taux d'intérêt et de l'écoulement du temps. Lorsque nous utilisons une stratégie avec une duration constante à chaque rebalancement, nous risquons de se trouver dans des situations de redondance ou de déficit.

Le principal inconvénient reste l'hypothèse de changements parallèles dans la structure des taux. Pour lever partiellement cet inconvénient, nous pouvons utiliser le deuxième terme du développement limite de la fonction $P = f(r_t)$. Nous introduisons alors la notion de convexité.

- La convexité de l'actif

La sensibilité permet de traiter de la variation du cours des obligations par rapport aux petites variations parallèles de la courbe des taux. Mais lorsque la variation des taux est forte, l'erreur obtenue en mesurant la variation relative du cours par la droite tangente devient importante.
Nous faisons donc recours au développement de Taylor du second ordre pour améliorer notre estimation de la variation de la valeur VA :

$$\Delta VA \approx VA' \times \Delta r + \frac{1}{2!} VA'' \times \Delta r^2$$

Formule où VA' et VA'' représentent respectivement la dérivée première et la dérivée seconde de $VA(r)$. En divisant par VA, Nous obtenons :

$$\frac{\Delta VA}{VA} \approx \frac{VA'}{VA} \times \Delta r + \frac{1}{2!} \frac{VA''}{VA} \times \Delta r^2$$

L'expression de la convexité est la suivante :

$$convexité = \frac{VA''}{VA} = \frac{1}{VA \times (1+r)^2} \times \left(\sum_{i=1}^{n} t_i \times (t_i + 1) \times \frac{F_{t_i}}{(1+r)^{t_i}} \right)$$

Avec cette définition, l'approximation de la valeur actuelle du cours de l'obligation est la suivante:

$$\Delta VA \approx VA \times \left[sensibilité \times \Delta r + \frac{convexité}{2} \times (\Delta r)^2 \right]$$

Comme la duration, la convexité est compatible avec l'analyse de la structure par terme des taux. Nous pouvons écrire :

$$convexité = \frac{1}{VA} \times \left(\sum_{i=1}^{n} t_i \times (t_i + 1) \times \frac{F_{t_i}}{(1+r_{t_i})^{t_i+2}} \right)$$

Comme pour la duration cette extension n'est valide que pour des variations parallèles de la courbe des taux.

Lorsque les flux de trésorerie sont tous fixes et positifs, la convexité est positive. Quel que soit le sens de la variation du taux d'actualisation, le terme $\frac{convexité}{2} \times (\Delta r)^2$ intervient positivement dans la variation du cours. Il est donc, habituel de considérer qu'une convexité supplémentaire liée à la dispersion des flux de trésorerie est « la bienvenue ».

L'insuffisance de l'indice de « duration », comme mesure de sensibilité, dans le cas systématique où les flux sont sensibles aux variations de taux a été la motivation majeure du développement d'autres approches (dont les modèles basés sur le surplus et les techniques de simulations). Le but est de fournir un indice de « sensibilité » plus adapté au cas des flux sensibles aux variations de taux.

II- Modèles basés sur le surplus

Ces modèles sont basés sur la notion du surplus S qui est souvent défini comme étant la différence entre la valeur de marché des actifs et la valeur actuelle du passif. Dans la suite, nous illustrons ces modèles via trois approches : celle de Kim et Santomero [1988], Sharpe et Tint [1990] et Leibowitz [1992]. Une quatrième approche permettant de lier simultanément les notions du surplus et de la duration sera également présentée.

II-1 Modèle de Kim et Santomero [1988]

Ce modèle donne plus d'importance aux conditions spécifiques à l'entreprise concernée. Celles-ci sont représentées par un ratio d'effet de levier : « *surplus/valeur de marché initiale de l'actif* ».

Il s'applique aux portefeuilles constitués de plusieurs classes d'actifs. Le but du modèle est de minimiser le risque de perte du surplus (mesuré par la variance de la rentabilité du surplus) sous contraintes de rentabilité et de poids des actifs.

Formulation théorique

La notion de rentabilité du surplus utilisée est la variation relative du surplus sur la période (année par exemple) :

$$R_S = \frac{S_1 - S_0}{S_0}$$

avec le surplus initial S_0 et S_1 le surplus final.

Nous considérons un portefeuille avec M classes d'actifs risqués et un passif évalué en valeur actuelle. Le vecteur des rentabilités attendues est : $R = [R_A, R_P]_{(M+1,1)}$

$(R_A)_{M\times 1}$ représente le vecteur rentabilités des actifs.

R_P la variation relative de la valeur actuelle du passif.

Le vecteur des poids est noté : $w = [w_A, w_P]$

Le poids de l'actif i est mesuré par la valeur de marché initiale de cet actif divisée par la valeur initiale du surplus.

$w_A^i = \frac{A_0^i}{S_0}$ poids de l'actif i.

$w_P = -\frac{P_0}{S_0}$ poids du passif.

Nous vérifions que la somme des poids est égale a un.

La rentabilité du surplus est donc :

$$R_S = w'_A R_A + w_P R_P$$

Ce modèle dépend d'un ratio $k_0 = \frac{S_0}{A_0}$ appelé ratio de solvabilité initial.

Nous vérifions que $k_0 = 1 - \frac{1}{rf_0}$ avec rf_0 le ratio de financement initial (valeur de marché des actifs divisée par la valeur actuelle du passif). Ce modèle permet d'obtenir une frontière efficiente en utilisant la même démarche que Markowitz [1952] : minimiser la variance de la rentabilité du surplus σ_s^2 pour un niveau de rentabilité du surplus donné avec une contrainte sur les poids (somme des poids égale a un).

Notations

Dans leurs travaux, Kim et Santomero [1988] utilisent les notations suivantes :

$$A = e' V_A^{-1} R_A \qquad B = R'_A V_A^{-1} R_A > 0$$
$$C = e' V_A^{-1} e > 0 \qquad D = BC - A^2 > 0$$
$$\mathrm{F}_0 = R'_A V_A^{-1} V_{AP} \qquad \mathrm{F}_1 = e' V_A^{-1} V_{AP}$$
$$F_2 = V'_{AP} V_A^{-1} V_{AP} > 0$$

Avec :
$e' = (1,...,1)$
V_A la matrice covariance des actifs
V_{AP} le vecteur covariance entre les actifs et le passif.

Equation analytique de la frontière efficiente

Le couple solution du problème vérifie la relation suivante, cette relation est l'équation d'une hyperbole dans le repère écart-type du surplus/rentabilité du surplus (cf. graphique 31) :

$$R_S = R_{SM} \pm \frac{1}{C}\left[DC\left(\sigma_S^2 - \sigma_{SM}^2\right)\right]^{1/2}$$

R_{SM} et σ_{SM} sont les caractéristiques du portefeuille de variance minimale (Coordonnées du sommet de l'hyperbole).

$$R_{SM} = \frac{1}{k_0}\frac{A}{C} + \left(\frac{1}{C}\right)\left(\frac{1}{k_0} - 1\right)\left(CF_0 - AF_1\right) - \left(\frac{1}{k_0} - 1\right)R_P$$
$$R_{SM} = \left(1 - \frac{1}{k_0}\right)^2 \left(Var\left(R_p\right) - F_2\right) + \left(\frac{1}{C}\right)\left(\frac{1}{k_0} - \left(\frac{1}{k_0} - 1\right)F_1\right)$$

La partie supérieure de l'hyperbole représente une frontière efficiente qui dépend du niveau du ratio k_0 (solvabilité de l'entreprise).

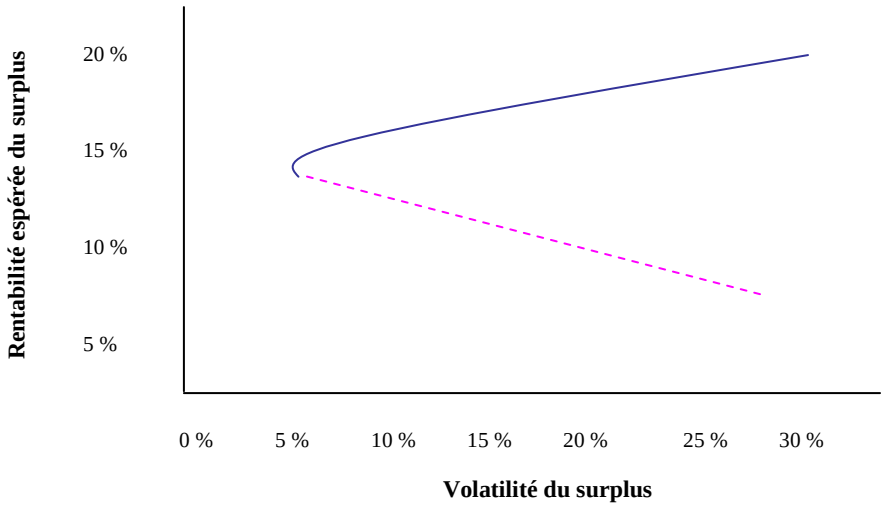


Fig. 31 : *Frontière efficiente (Kim&Santomero)*

Contrainte de déficit

Nous supposons que le vecteur des rentabilités $R = [\, R_A, R_P \,]_{(M+1,1)}$, est gaussien (la variable rentabilité du surplus suit donc une loi normale). Nous introduisons la contrainte de déficit suivante : « La probabilité pour que la rentabilité du surplus soit inférieure à un certain seuil ne doit pas dépasser une probabilité donnée ». Ce qui revient à limiter la probabilité de perte d'une partie du surplus initial.

Cette contrainte est représentée par une droite de déficit dans le repère (écart-type de la rentabilité du surplus, rentabilité du surplus espérée). L'intersection entre cette droite et la frontière efficiente donne le portefeuille efficient vérifiant la contrainte (cf. graphique 32).

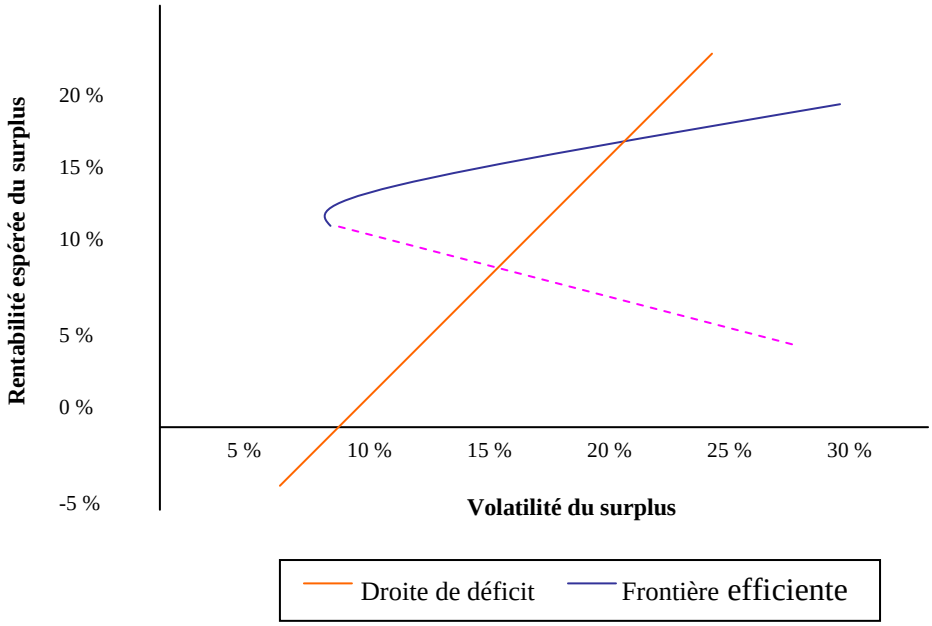


Fig. 32 : *Frontière efficiente et contrainte de déficit (Kim&Santomero)*

Nous pouvons aussi, en utilisant les techniques de Roy [1952], chercher le portefeuille efficient qui minimise la probabilité de perdre un pourcentage du surplus initial. Cette approche a deux avantages :

- La prise en compte de la corrélation entre la valeur de marché de l'actif et la valeur actuelle du passif
- L'introduction d'une relation entre la solvabilité de l'entreprise et l'allocation optimale d'actifs.

II-2 Modèle de Sharpe et Tint [1990]

Ce modèle s'applique aussi à un portefeuille composé de plusieurs classes d'actifs. Le passif est évalué en valeur actuelle et l'actif en valeur de marché. Il est basé sur la notion de surplus S défini comme étant la différence entre la valeur de marché des actifs et la valeur actuelle du passif. Le but du modèle est de minimiser le risque de perte du surplus (mesuré par la variance de la « rentabilité du surplus ») pour un niveau de « rentabilité du surplus » donné et sous un ensemble de contraintes (même démarche que Markowitz [1952]).

La notion de rentabilité du surplus utilisée est la variation du surplus divisée par la valeur initiale des actifs :

$$R_S = \frac{S_1 - S_0}{A_0}$$

L'ensemble des portefeuilles vérifiant ce programme d'optimisation pour différents niveaux de rentabilité espérée, décrit *une hyperbole* dans le repère (risque mesuré par l'écart type de la rentabilité du surplus, rentabilité espérée du surplus). La partie supérieure de cette hyperbole représente *la frontière efficiente*.

Formulation théorique

Le portefeuille est composé de n actifs risqués de rentabilité $R_i\,(i=1,...,n)$

La stratégie d'investissement est donnée par le choix du portefeuille $x=(x_i)_{i=1,...,n}$ avec $\sum_{i=1}^{n} x_i = 1$

$R_A(x)=\sum_{i=1}^{n} x_i R_i$: la rentabilité du portefeuille des actifs.

R_p : la variation relative de la valeur actuelle du passif.

rf_0 : ratio de financement initial

$R_s(x)=R_A(x)-\frac{1}{rf_0}R_P$: la rentabilité du surplus.

$V=\left[cov(R_i,R_j)\right]$ la matrice de variance covariance supposée régulière,

$\delta=(cov(R_i,R_P))$ le vecteur des covariances entre les actifs et le passif,

$\mu=(E(R_i))$ le vecteur des rentabilités espérées des actifs,

$e=(1,...,1)$: le vecteur unité transposé

x^* : Portefeuille optimal,

x^{min} : Portefeuille à variance minimale,

Les portefeuilles efficients

Le portefeuille x^* solution du problème d'optimisation est donné par la relation suivante :
$x^* = x^{min} + \lambda z^*$

λ dépend du niveau de rentabilité du surplus.
$x^{min} = \frac{V^{-1}e}{e'V^{-1}e} + \frac{1}{rf_0}\left(V^{-1}\delta - \frac{e'V^{-1}\delta}{e'V^{-1}e}V^{-1}e\right)$ représente le portefeuille a variance minimale.
$z^* = V^{-1}\mu - \frac{e'V^{-1}\mu}{e'V^{-1}e}V^{-1}e$

La frontière efficiente

En remplaçant le portefeuille efficient par son expression, nous obtiendrons les deux équations suivantes :

$$Var\left(R_s\left(x^*\right)\right) = \lambda^2 Var\left(R_A\left(z^*\right)\right) + Var\left(R_s\left(x^{min}\right)\right)$$
$$E\left(R_s\left(x^*\right)\right) = \lambda E\left(R_A\left(z^*\right)\right) + E\left(R_S\left(x^{min}\right)\right)$$

En éliminant le λ dans les deux équations, nous obtenons l'équation d'une hyperbole. La partie supérieure de cette courbe représente la frontière efficiente dans le repère (volatilité du surplus, rentabilité attendue du surplus).

Contrainte de déficit

« La probabilité pour que la rentabilité du surplus soit inférieure à un certain seuil (u) ne doit pas dépasser une certaine probabilité (p) ».
Si nous supposons que le vecteur des rentabilités est gaussien, la variable aléatoire rentabilité du surplus suit une loi normale. La contrainte de déficit se traduit par l'équation de droite suivante dans le plan (volatilité du surplus, rentabilité espérée du surplus) :

$$E(R_S) = q_P^{N(0,1)} * \sigma_S + u$$

$q_P^{N(0,1)}$: représente le quantile de la loi normale centrée réduite associé à la probabilité 1-*p*.

L'intersection entre la droite de déficit et la frontière efficiente donne le ***portefeuille efficient de rentabilité maximale*** *et* ***vérifiant cette contrainte***.

Remarque: Une contrainte de déficit sur le ratio de financement

Le ratio de financement à l'instant final (grandeur inconnue) est donné par la relation suivante :

$$rf_1(x) = rf_0 \frac{1 + R_A(x)}{1 + R_P}$$

La contrainte de déficit est la suivante: «la probabilité pour que le ratio de financement final soit inférieur à 100 % ne doit pas dépasser une certaine probabilité *p*».

Cette contrainte se traduit par l'équation de droite suivante :

$$E\left(R_S\right) = q_P^{N(0,1)} * \sigma_S + \left(\frac{1}{rf_0} - 1\right)$$

Les courbes obtenues dans ce cadre (cf. graphique 33) auront ainsi la même allure obtenue dans le graphique précédent (avec le modèle de *Kim&Santomero*).

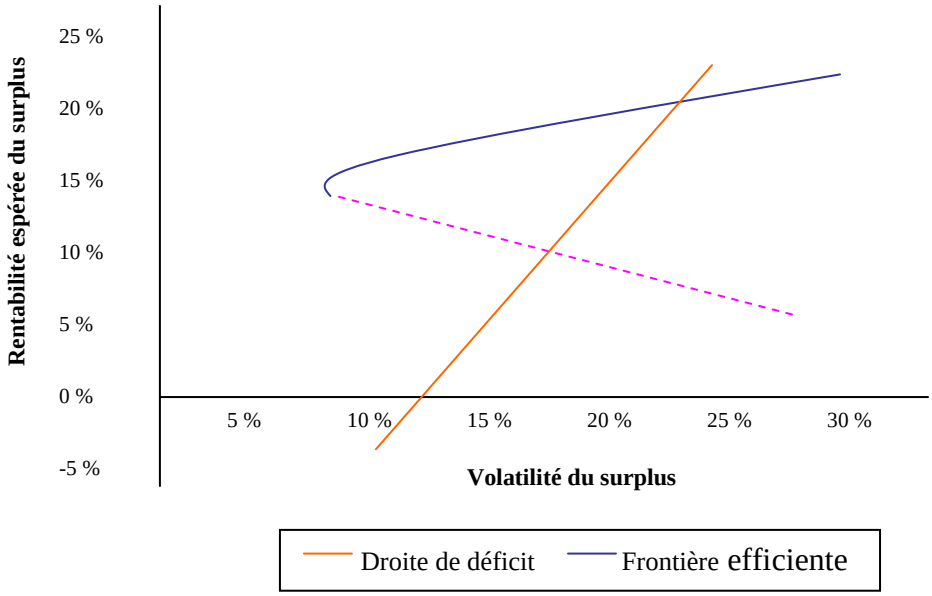


Fig. 33 : *Frontière efficiente et contrainte de déficit (Sharpe &Tint)*

II-3 Modèle de Leibowitz [1992]

Ce modèle s'applique à un portefeuille composé de deux actifs: actions et obligations. Il consiste à déterminer d'une part le pourcentage d'actions et d'autre part la « duration » de la composante obligataire qui permettent de maximiser la « rentabilité du surplus » et donc la rentabilité du portefeuille tout en respectant des contraintes sur l'actif et le passif. Ces contraintes sont modélisées en probabilité « d'insuffisance » et concernent des grandeurs et des rentabilités exprimées en valeur de marché ou valeur actuelle.

a - Hypothèses et notations :

R_A : rentabilité des actions (suit une loi normale).
$\overline{R}_A$: rentabilité attendue des actions.
σ_A : volatilité des actions, écart-type de la rentabilité RA.
R_O : rentabilité des obligations (suit une loi normale).
$\overline{R}_O$: rentabilité attendue des obligations.
σ_O : volatilité des obligations, variable endogène au modèle.
ρ : corrélation (R_A , R_O)
R_{pf} : rentabilité du portefeuille (suit une loi normale de moyenne $\overline{R}_{pf}$).
$R_{pf} = \alpha R_A + (1-\alpha) R_O$
α : pourcentage d'actions.
σ_{pf} : volatilité du portefeuille. $\sigma_{pf}^2 = \alpha^2 \sigma_a^2 + (1-\alpha)^2 \sigma_0^2 + 2\alpha(1-\alpha)\sigma_A \sigma_O \rho$

R_P : Variation relative des valeurs actuelles du passif.
rf_0 : Ratio de financement initial (valeur de marche initiale de l'actif divisée par la valeur actuelle initiale du passif).
D_0 : duration des obligations.

La volatilité de la composante obligataire est supposée proportionnelle à sa duration.

$$\sigma_0 = D_0 \sigma_{taux1an}$$

$\sigma_{taux1an}$: écart-type du taux d'intérêt pour les emprunts 1 an.

Les obligations de toutes les maturités fournissent le même rendement.

La rentabilité du surplus : mesure *de risque actif / passif*

La rentabilité du surplus est représentée dans ce modèle par le rapport entre la variation du surplus et la valeur actuelle initiale du passif.

$$R_S = \frac{S_1 - S_0}{P_0}$$

S_0 le surplus initial, S_1 le surplus final et P_0 la valeur actuelle initiale du passif. En décomposant l'expression du surplus, la rentabilité du surplus devient : $R_S = rf_0 * R_{pf} - R_P$

b - La contrainte sur l'actif

« La probabilité pour que la rentabilité du portefeuille d'actifs soit inférieure à un certain seuil ne doit pas dépasser une probabilité donnée »

$$P(R_{pf} < u') < p'$$

La probabilité d'insuffisance p' et le seuil critique u' sont fixés arbitrairement selon la tolérance au risque de l'investisseur. Cette contrainte dans le plan risque/rentabilité espérée est caractérisée par une droite de « déficit ». (Tous les portefeuilles au dessus de cette droite vérifient cette contrainte).

$$\overline{R}_{pf} = u' + \sigma_{pf} * q_{p'}^{N(0,1)}$$

c - La contrainte sur le surplus

« La probabilité pour que la rentabilité du surplus soit inférieure a un certain seuil (u) ne doit pas dépasser une probabilité donnée (p) ».

$P(R_S < u) < p$

Nous supposons que R_{pf} et R_P sont normalement distribuées, et que toute combinaison linéaire de ces deux variables aléatoires est aussi normalement distribuée ($R_{pf} \approx N(\overline{R}_{pf}, \sigma_{pf})$ et $R_P \approx N(\overline{R}_P, \sigma_P)$).

Nous avons donc :

$$R_S \approx N(\overline{R}_S, \sigma_S)$$

De la même façon que la contrainte de « déficit » sur l'actif, nous déduisons l'inéquation suivante :

$$u + \sigma_S q_P^{N(0,1)} < \overline{R}_S$$

Démarche théorique :

La frontière de la courbe qui décrit cette contrainte est caractérisée par :

$$u + \sigma_S q_P^{N(0,1)} = \overline{R}_S$$

$$\overline{R}_S = rf * \overline{R}_{pf} - \overline{R}_P$$
$$\sigma_S^2 = rf^2 * \sigma_{pf}^2 + \sigma_P^2 - 2 * rf * \sigma_{pf} * \sigma_P * Corr\left(R_{pf}, R_p\right)$$

L'équation de la frontière devient après substitution des expressions ci-dessus :

$$\left[rf^2(1-\alpha)^2\right]\sigma_0^2 + \left[2rf^2\alpha(1-\alpha)\sigma_A\rho - 2rf(1-\alpha)\sigma_P\right]\sigma_0$$
$$-\left[\left(\frac{rf\left(\alpha\overline{R}_A + (1-\alpha)\overline{R}_O\right) - \overline{R}_P - u}{q_P^{N(0,1)}}\right)^2 - \sigma_P^2 - rf^2\alpha^2\sigma_A^2 + 2rf\alpha\sigma_A\sigma_P\rho\right] = 0$$

Les seules inconnues restantes sont le pourcentage d'action α et la volatilité de l'obligation à détenir dans le portefeuille σ_0 .

Pour un α donné, cela revient à résoudre une équation du second degré en σ_0. La condition de positivité sur le discriminant de cette équation permet d'obtenir un ensemble de valeurs de α à partir desquelles nous calculons les σ_0 correspondants. Grâce aux valeurs de ces deux variables, nous reconstituons un ensemble de portefeuilles de coordonnées (σ_{pf} ,$\overline{R}_{pf}$) dans le repère risque/rentabilité dont les surplus associés vérifient la contrainte de déficit. La contrainte sur le surplus est représentée par une courbe convexe en forme d'œuf.

d - Portefeuille optimal vérifiant les deux contraintes

La partie supérieure à la droite de déficit et à l'intérieur de « l'œuf » représente tous les portefeuilles qui vérifient les deux contraintes. Le portefeuille intersection de la droite et de « l'œuf » de surplus, de caractéristiques $\overline{R}_{pf}^*$ et σ_{pf}^* , correspond au portefeuille de rentabilité espérée la plus grande qui remplit les deux demandes de « déficit ».
La part optimale d'actions à détenir en portefeuille :

$$\alpha^* = \frac{\overline{R}_{pf}^* - \overline{R}_O}{\overline{R}_A - \overline{R}_O}$$

La volatilité optimale des obligations à détenir en portefeuille :

$$\sigma_O^* = \frac{\sqrt{\left(\sigma_{pf}^*\right)^2 + \left(\alpha^*\right)^2\sigma_A^2\left(\rho^2 - 1\right)} - \alpha^*\sigma_A\rho}{1-\alpha^*}$$

e- L'impact des changements des paramètres

Variation du seuil de rentabilité de l'actif :

Les changements du seuil de rentabilité de l'actif entraînent des changements parallèles de la droite de déficit comme le montre le graphique suivant :

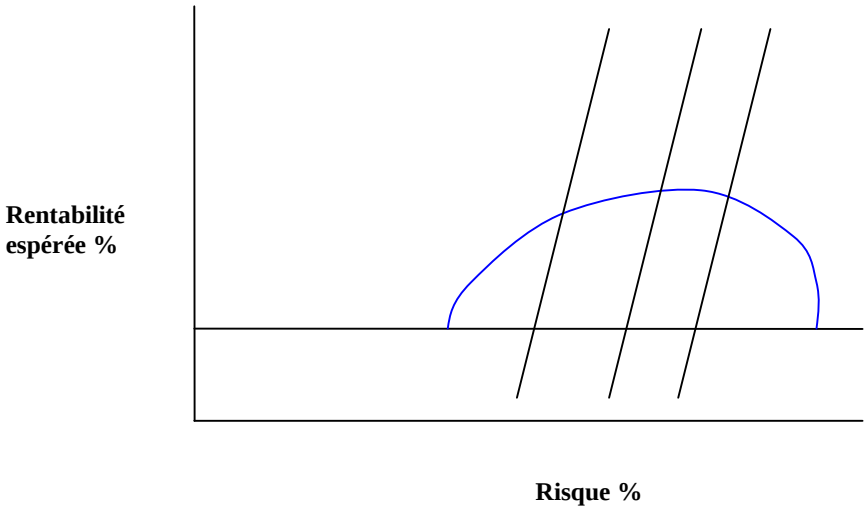


Fig. 34 : *Exemple de l'impact de la variation du seuil de rentabilité de l'actif*

Variation de la probabilité de déficit de l'actif :

La pente de la droite de «déficit» croit lorsque la probabilité p' diminue, alors que le point d'intersection de cette droite avec l'axe des ordonnées reste fixe (cf. graphique 35).

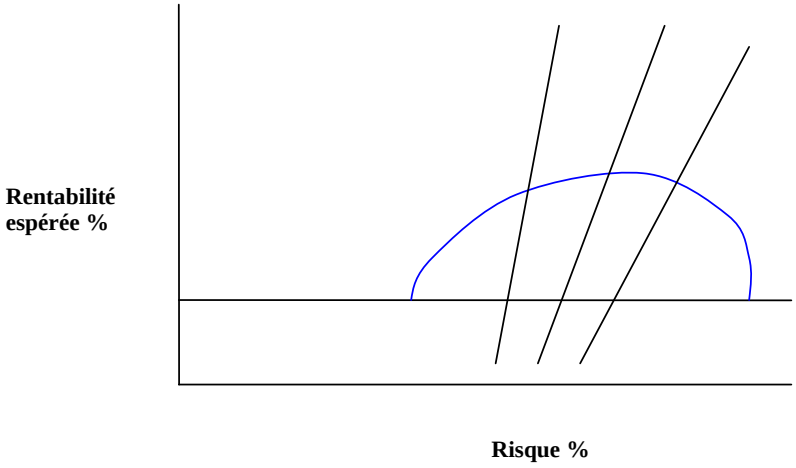


Fig. 35 : *Exemple de l'impact de la variation de la probabilité de déficit de l'actif*

Variation du seuil de rentabilité du surplus :

Lorsque le seuil de rentabilité du surplus augmente, la taille de l'œuf diminue et la rentabilité du portefeuille optimal aussi (pour une contrainte fixe sur l'actif). Cela est illustré dans le graphique suivant :

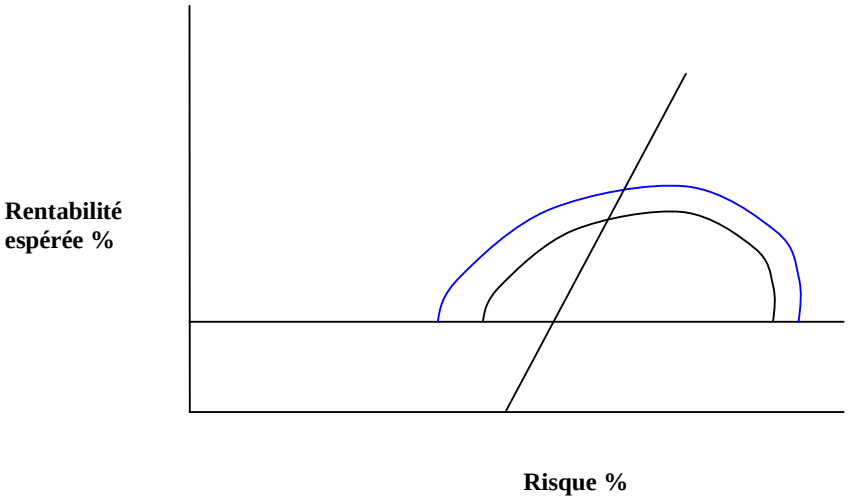


Fig. 36 : *Exemple de l'impact de la variation du seuil de rentabilité du surplus*

Variation de la probabilité de déficit du surplus :

L'œuf de déficit varie de la même façon que précédemment.

Variation du ratio de financement initial :

Plus le ratio de financement initial diminue, plus l'œuf de déficit rétrécit et se déplace vers la droite (cf. graphique 37).

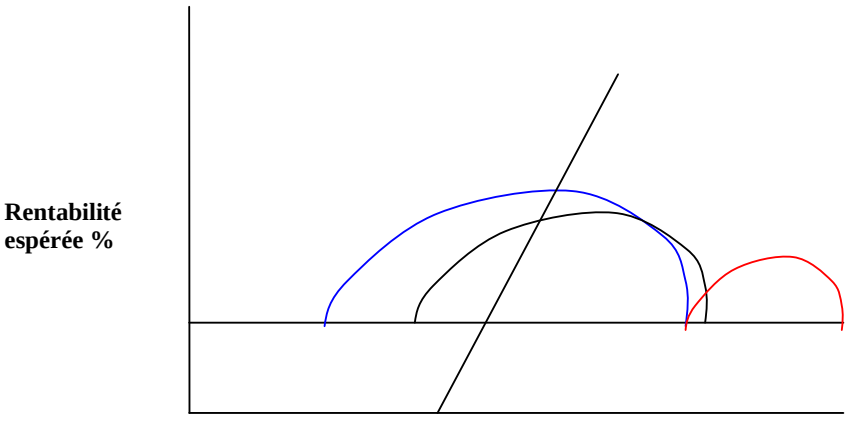


Fig. 37 : *Exemple de l'impact de la variation du ratio de financement initial*

f- Contrainte sur la rentabilité relative par rapport à un benchmark

« La probabilité pour que la rentabilité du portefeuille d'actifs soit inférieure d'un certain seuil à la rentabilité du benchmark ne doit pas dépasser une certaine probabilité ».

$$P\left(R_{rel} < u''\right) = p''$$

R_{rel} : La rentabilité relative, c'est l'écart entre la rentabilité du portefeuille d'actifs et la rentabilité du benchmark. Le benchmark considéré est composé de deux actifs : actions et obligations. Le pourcentage en actions et la duration de la composante obligataire du benchmark sont connus.

Les rentabilités du portefeuille d'actifs et du benchmark sont données par :

$$R_{ptf} = aR_A + (1-a)R_{O_1}$$
$$R_{ben} = bR_A + (1-b)R_{O_2}$$

R_A : La rentabilité de l'action qui est la même pour le portefeuille d'actifs et le benchmark,

R_{O_1} : La rentabilité de la composante obligataire du portefeuille,

R_{O_2} : La rentabilité de la composante obligataire du benchmark,

a : le pourcentage d'actions dans le portefeuille, à déterminer.

b : le pourcentage d'actions dans le benchmark, connu.

La rentabilité relative a pour expression :

$$R_{rel} = (a-b)R_A + (1-a)R_{O_1} + (1-b)R_{O_2}$$

La rentabilité relative espérée est donc :

$$\overline{R}_{rel} = (a-b)\overline{R}_A + (1-a)\overline{R}_{O_1} + (1-b)\overline{R}_{O_2}$$

Nous supposons que les deux obligations ont la même rentabilité espérée ($\overline{R}_{O_1} = \overline{R}_{O_2}$):

$$\overline{R}_{rel} = (a-b)\left(\overline{R}_A - \overline{R}_{O_1}\right)$$

C'est donc le produit de « l'excès » d'actions ($a-b$) et de la prime de risque.

La variance de la rentabilité relative est donnée par :

$$\sigma_{rel}^2 = E\left[\left(R_{rel} - \overline{R}_{rel}\right)^2\right]$$

Avec l'hypothèse de normalité, et en posant :

$$corr(A,O_1) = corr(A,O_2) = \rho$$

nous obtenons :

$$\sigma_{rel}^2 = (a-b)^2\sigma_A^2(1-\rho^2) + \left[(a-b)\sigma_A\rho + (1-a)\sigma_{O1} - (1-b)\sigma_{O_2}\right]^2$$

Cette relation donne une formule usuelle pour la variance de la rentabilité relative et permet de trouver des portefeuilles qui rencontrent la contrainte de déficit.

$$P\left(R_{rel} < u''\right) = p''$$

Cette contrainte se traduit par la relation suivante (hypothèse de normalité) :

$$\overline{R}_{rel} = u'' + q_{p''}^{N(0,1)}\sigma_{rel}$$

Les seules inconnues sont le pourcentage d'actions a et la volatilité de l'obligation à détenir dans le portefeuille σ_{O_1}. Pour un a donné, cela revient à résoudre une équation du second

degré en σ_{O_1} . La condition de positivité sur le discriminant du trinôme permet d'obtenir un ensemble de valeurs de « a » à partir duquel nous calculons les σ_{O_1} correspondants. Les valeurs de ces deux variables permettent de reconstituer un ensemble de portefeuilles de coordonnées (σ_{pf} ,$\overline{R}_{pf}$) dans le repère risque/rentabilité espérée vérifiant la contrainte de déficit. L'ensemble des portefeuilles vérifiant la contrainte sur rentabilité relative décrit une courbe convexe en forme d'œuf comme le montre le graphique suivant :

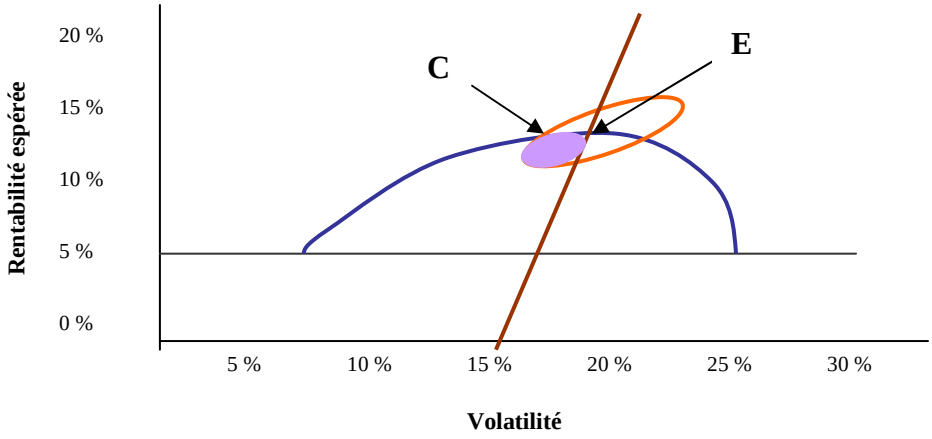


Fig. 38 : *Représentation graphique des portefeuilles vérifiant la contrainte sur la rentabilité relative*

L'ensemble des portefeuilles optimaux respectant les trois contraintes de Leibowitz se situe dans la zone grise. Les portefeuilles efficients se situent alors sur l'arc [CE]. Le portefeuille E correspond au portefeuille efficient répondant à la recherche d'un rendement optimal. Le point C indique un portefeuille efficient qui a un risque inférieur. Puisque les prévisions de l'investisseur ne sont jamais sures à 100 %, il est raisonnable de préférer un point compris entre E et C à la place d'un point extrême.

g- Limite du modèle et test de robustesse

Les limites de ce modèle résident dans le fait qu'il est basé sur des hypothèses très fortes :
- la volatilité de la composante obligataire est proportionnelle à sa « duration ».
- les obligations de toutes les maturités fournissent le même rendement.
- les rentabilités sont normalement distribuées.

Ce modèle économique ne respecte pas les contraintes règlementaires, techniques, et comptables. Il surestime donc la capacité de prise de risque. Pour tester la robustesse du modèle, nous faisons varier les hypothèses sur les entrées (corrélation, volatilité et rentabilités attendues) et nous mesurons l'impact de ses variations sur le pourcentage d'actions et la duration de la composante obligataire.

h- Une variante du modèle de Leibowitz

Pour mesurer le risque actif/passif, nous remplaçons la rentabilité du surplus par une mesure plus concrète : la « rentabilité du ratio de financement ». C'est la variation relative du ratio de financement sur la période considérée. Si nous notons rf_0 et rf_1 les ratios de financement initial et final, la « rentabilité du ratio de financement » notée RRF est donnée par :

$$RRF = \frac{rf_1 - rf_0}{rf_0}$$

Cette mesure dépend uniquement de la rentabilité de l'actif et de la variation relative de la valeur actuelle du passif et ne dépend pas du ratio de financement initial.

Nous notons :

A_0 et A_1 les valeurs de marché de l'actif en début et fin de période,

P_0 et P_1 les valeurs actuelles du passif en début et fin de période,

$R_{pf} = \frac{A_1 - A_0}{A_0}$ et $R_p = \frac{P_1 - P_0}{P_0}$

Nous obtenons :

$$RRF = \frac{R_{pf} - R_P}{1 + R_P}$$

Nous pouvons intégrer la rentabilité du ratio de financement dans l'analyse d'allocation d'actif sous forme de contrainte de déficit : « la probabilité pour que la rentabilité du ratio de financement soit inférieure a un certain seuil ne doit pas dépasser une probabilité donnée ». Le seuil de déficit dépend contrairement à la RRF du ratio de financement initial. Nous modifions le modèle de Leibowitz en remplaçant la contrainte sur la rentabilité du surplus par la contrainte sur la RRF .

$$P(RRF < u) < p$$

Cette contrainte est représentée par une courbe convexe.

II-4 Duration du Surplus

Le surplus est la différence entre la valeur de marche de l'actif et la valeur actuelle du passif.

Nous introduisons les notations suivantes :

- A : valeur de marché de l'actif
- P : valeur actuelle du passif
- $S = A - P$: surplus
- $rf = \frac{A}{P}$: le ratio de financement
- r : le taux d'intérêt
- $D_A = -\frac{1}{A}\frac{\Delta A}{\Delta r}(1+r)$ la duration de l'actif
- $D_P = -\frac{1}{P}\frac{\Delta P}{\Delta r}(1+r)$ la duration du passif

Nous définissons la « duration » du surplus de la même façon :

$$D_S = -\frac{1}{S}\frac{\Delta S}{\Delta r}(1+r)$$

En remplaçant S par son expression, nous obtenons la formule suivante :

$$SD_S = AD_A - PD_P$$

En divisant cette expression par le surplus S, nous obtenons :

$$D_S = D_P + \frac{rf}{rf - 1}(D_A - D_P)$$

Certaines remarques peuvent être avancées :

- Un adossement par la duration n'élimine pas le risque de taux sur le surplus. En effet, la formule précédente donne une duration du surplus égale à la duration du passif lorsque le ratio de financement est différent de 100 % : $D_S = D_A = D_P$.
- La duration de l'actif qui permet d'éliminer le risque de taux sur le surplus (duration du surplus nulle) est donnée par : $D_A^* = \frac{1}{rf} D_P$

Pour un ratio de financement supérieur à 100 %, l'actif obtenu est plus court que le passif.

- Lorsque le surplus est nul, la duration du surplus n'est pas définie.
- Nous pouvons représenter graphiquement la duration du surplus en fonction de la différence entre la duration de l'actif et la duration du passif.

La courbe obtenue est une droite qui traduit tous les niveaux de risque de taux sur le surplus en fonction de cette différence.

Pour un ratio de financement supérieur à 100 % (pente de la droite positive), nous obtenons la courbe suivante :

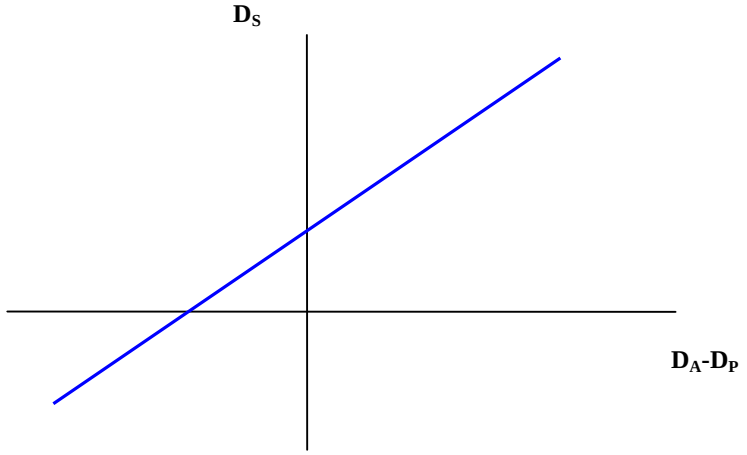


Fig. 39 : *Exemple de la variation de la duration du surplus en fonction de l'écart entre la duration de l'actif et celle du passif*

Notons que nous pouvons définir de la même façon la «convexité» du surplus notée C_S :

$$C_S = \frac{1}{2S} \frac{\Delta^2 S}{\Delta r^2}$$

Si nous notons respectivement C_A et C_P les convexités de l'actif et du passif, nous obtenons la formule suivante :

$$C_S = C_P + \frac{rf}{rf - 1}(C_A - C_P)$$

Conclusion du chapitre 1

L'aspect mono-périodique qui caractérise les différents modèles classiques de gestion actif-passif (que ce soit celles basées sur les techniques d'immunisation ou celles basées sur la notion du surplus) limite leur capacité à refléter de façon fiable les perspectives d'évolution des différentes variables financières, en particulier sur le long terme. Cela devient d'autant plus compliqué que les variables à considérer sont plus nombreuses et que la nécessité de prendre en compte les corrélations entre les différentes variables est plus importante.

Nous notons comme même que les modèles classiques de la gestion actif-passif, servant en particulier pour l'allocation stratégique d'actif, étaient de loin l'outil de référence sur le marché et ce depuis les années 1950. Les limites que présentent ces modèles ainsi que le développement des supports informatiques au cours des années 1980 ont permis récemment de passer vers d'autres types de modèles : les modèles stochastiques d'ALM. Ces derniers seront l'objet du chapitre suivant.

Chapitre 2 : Allocation stratégique d'actifs et stratégie *Fixed-Mix*

Les modèles stochastiques d'ALM utilisent les techniques de simulation stochastique (*Monte Carlo*) pour modéliser l'évolution des différents éléments, que ce soit au niveau des actifs financiers et des engagements, ou au niveau des variables de marché et des variables démographiques. Nous pourrons ainsi estimer les lois de probabilité associées aux résultats du fonds de retraite sur le long terme. Un passage vers le cadre multi-périodique est effectué. Dans un cadre d'une allocation stratégique stochastique, la revue de certains récents travaux peut être utile.

Frauendorfer [2007] montre comment un critère moyenne-variance peut être appliqué dans un cadre multi-périodique afin d'obtenir les portefeuilles efficients d'une gestion actif-passif. Le modèle d'optimisation tient compte des coûts de transaction et des volatilités stochastiques aussi bien des actifs que du passif. De plus, un outil général permettant la projection des engagements du fonds de pension ainsi que la projection des rendements des actifs est présenté. Dans une étape suivante, la dynamique de la structure par maturité des engagements (*liability maturity structure*) est modélisée comme un indice personnalisé (*customized index*) dont la volatilité et la corrélation avec les rendements de l'actif deviennent une composante intégrale de l'approche de changement de régime appliquée. Les résultats numériques illustrent la diversification des actifs et la relation entre la variabilité du rendement et la dynamique des engagements.

Martellini [2006] considère le problème de sélection de portefeuille dans un cadre intertemporel, en présence de contraintes sur le passif. En utilisant la valeur des engagements comme un numéraire naturel (*natural numeraire*), il trouve que la solution à ce problème induit un théorème de séparation en trois fonds. Cela constitue une justification à certaines pratiques appelées *Liability-Driven-Investment* (LDI), offertes par différentes banques d'investissement et des sociétés de gestion de portefeuille. La LDI se base sur l'investissement en deux types de fonds (en plus de l'actif sans risque) : le portefeuille de performance et le portefeuille de couverture du passif. Autrement, cette stratégie consiste à séparer le portefeuille en deux parties, dont l'une recourt à l'immunisation contingente et dont l'autre cherche le rendement absolu. Pour cette seconde partie le contrôle de risque passe impérativement par les méthodes de tests de stress capables d'évaluer les risques extrêmes.

Waring [2004] fait la revue, tout en les mettant à jour, des différentes techniques de détermination des frontières efficientes de surplus ainsi que l'allocation d'actif du surplus (*surplus asset allocation*). L'actualisation de ces techniques se fait à travers la prise en compte des caractéristiques systématiques (béta) et non systématiques (alpha). Il développe ainsi une vision économique des engagements, en termes d'alpha et de béta. Cela nous donne une mesure des engagements qui est plus intéressante pour résoudre le problème d'allocation d'actif et ce par rapport à ce qui est fourni dans les approches standards. Avec ces outils nous pouvons améliorer le contrôle des risques pour les fonds de pension, à travers le contrôle de la couverture du passif par les actifs. En plus, cette vision favorise, de façon appropriée, le recours à l'alpha et plus généralement à la mesure du risque dynamique provenant de la gestion active.

Munk et al. [2004] présentent la stratégie optimale d'allocation d'actifs pour un investisseur qui peut investir en *cash* (monétaire), en obligations nominales et en actions (ou indices d'actions). Le modèle suppose le retour à la moyenne des rendements des actions et des taux de rendement réels aléatoires. Le modèle de marché est calibré pour le cas des données

historiques des Etats-Unis (actions, obligations et inflation). En plus, afin d'illustrer les allocations d'actifs optimales, Munk et al. [2004] présentent une méthode de calibrage où les paramètres d'aversion au risque et les horizons d'investissement sont estimés de façon à avoir les recommandations optimales pour différents groupes d'investisseurs : « agressif », « modéré » et « conservatif » à différentes horizons d'investissement.

A ce niveau, nous nous proposons de distinguer deux sous-groupes de modèles d'ALM basés sur les techniques stochastiques. L'élément clé de distinction sera si oui ou non les poids des différents actifs reviennent périodiquement à ceux de l'allocation stratégique définie initialement (si oui, les modèles seront appelés modèles à poids constants ou stratégie *Fixed-Mix*).

- Pour le premier sous-groupe de modèles à poids constants et malgré les avancées réalisées avec ces techniques (surtout au niveau de l'implémentation informatique), l'aspect dynamique de l'allocation stratégique reste encore marginalisé. En fait, ces modèles permettent de comparer des allocations constantes dans le temps (statiques) indépendamment des opportunités liées aux évolutions inter-temporelles des marchés (cf. Merton [1990], Kouwenberg [2001], Infanger [2002], Dempster et al. [2003]). L'étude de l'allocation stratégique d'actifs dans le cadre des modèles à poids constants est l'objet de ce chapitre.

- Le deuxième sous-groupe de modèles, et le plus récent, est principalement inspiré de la théorie du choix de la consommation et de portefeuille développée par Merton [1971]. Il s'agit des modèles d'allocation dynamique. Par exemple, à partir de la définition de la fonction objectif pour l'investisseur, ces modèles permettent la détermination d'une trajectoire des poids des différents actifs jusqu'à la date d'échéance (l'ajustement des poids est fonction des évolutions projetées du marché et de la règle de gestion prédéfinie). L'allocation stratégique retenue sera l'allocation optimale d'actifs à la date initiale t_0. Le cadre d'utilisation de ces modèles récents se heurte au problème d'implémentation vu la complexité des outils mathématiques employés (cf. Hainaut et al. [2005], Hainaut et al. [2007], Rudolf et al. [2004] et Yen et al. [2003]).

Dans ce chapitre, nous considérons une approche d'allocation stratégique d'actifs basée sur la stratégie dite « à poids constants » ou *Fixed-Mix*. Nous proposons une modélisation du régime-type de retraite et étudions certains critères d'allocation stratégique d'actifs en fonction du type du régime (provisionné, partiellement provisionné, etc.) sont étudiés. Nous passons ensuite à l'illustration de la stratégie *Fixed-Mix* avec une application sur les réserves d'un régime de retraite partiellement provisionné. Les résultats obtenus sont discutés et différents tests de sensibilité sont mises en place : ces tests sont liés principalement à l'impact des hypothèses de rendement ou de corrélation retenues (le graphique 40 illustre les étapes générales nécessaires pour l'aboutissement d'un processus d'ALM stochastique).

Les difficultés rencontrées lors de la mise en place de la stratégie *Fixed-Mix*, dues essentiellement à la multiplicité du nombre de classes d'actifs considérées et au nombre de scénarios économiques simulés, nous ont mené à nous pencher sur les aspects d'optimisation numérique. Le point de départ est notre constat d'un temps de calcul significatif dû

simultanément à un nombre élevé de scénarios économiques générés et à un nombre d'allocations d'actifs testées également élevé.

Dans ce cadre, une présentation détaillée des travaux menés lors de la rédaction de l'article de Rullière et al. [2010] est effectuée. Un algorithme d'optimisation globale d'une fonction non convexe et bruitée est présenté. L'algorithme est construit après une étude de critères de compromis entre, d'une part, l'exploration de la fonction objectif en de nouveaux points (correspondant à des tests sur de nouvelles allocations d'actifs) et d'autre part l'amélioration de la connaissance de celle-ci, par l'augmentation du nombre de tirages en des points déjà explorés (correspondant à la génération de scénarios économiques supplémentaires pour les allocations d'actifs déjà testées). Une application numérique illustre la conformité du comportement de cet algorithme à celui prévu théoriquement et compare les résultats obtenus avec l'algorithme de Kiefer-Wolfowitz- Blum (cf. Blum [1954], Kiefer et Wolfowitz [1952]).

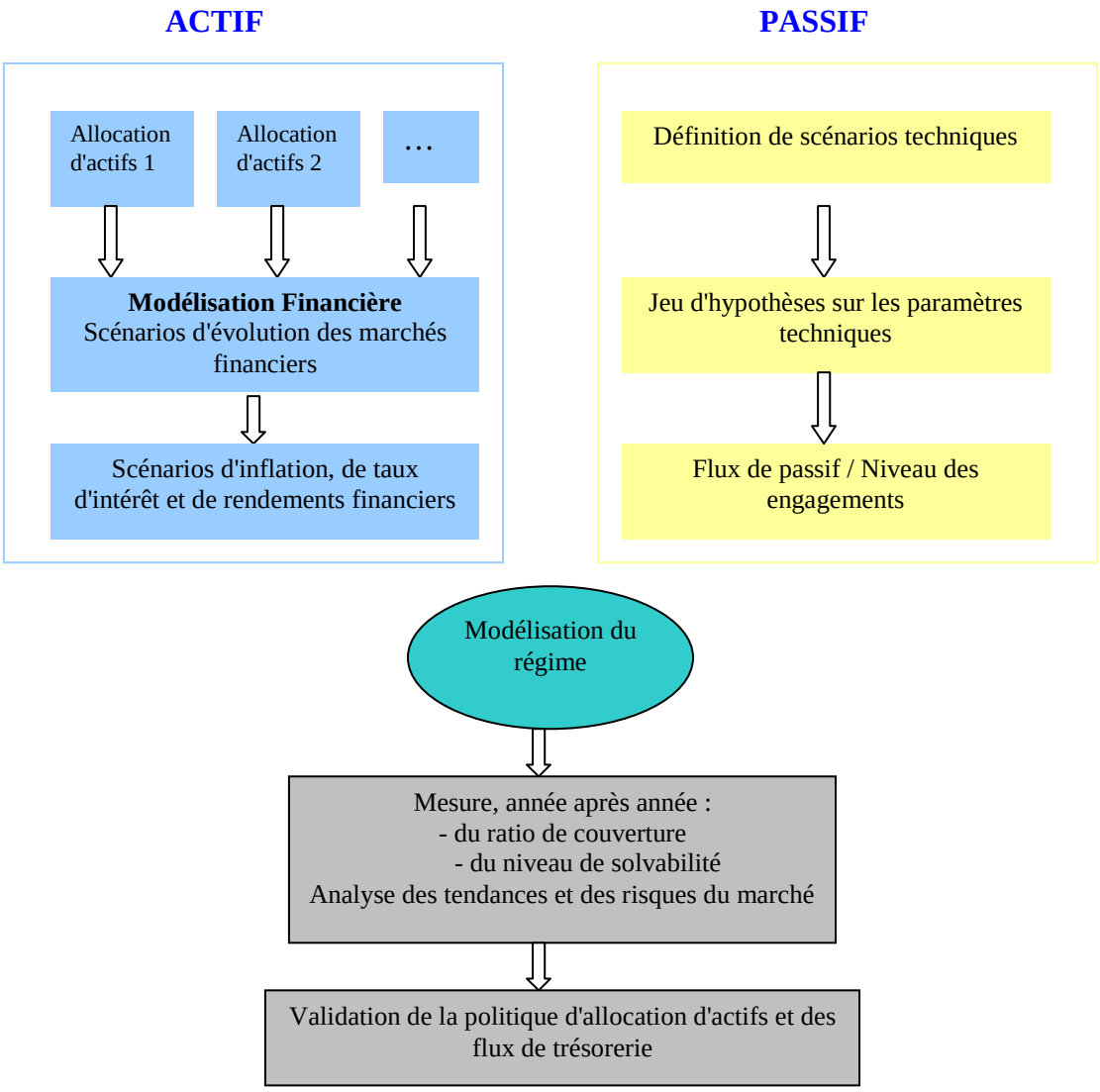


Fig. 40 : *Illustration du modèle de gestion actif-passif dans le cadre stochastique*

I- Etude de l'allocation d'actifs dans le cadre de la stratégie *Fixed-Mix*

I-1 Présentation

La gestion à poids constants est une gestion pour laquelle les poids reviennent périodiquement à ceux de l'allocation stratégique long-terme. Autrement dit, à chaque fin de période de simulation nous reconstituons le fonds géré (en effectuant les opérations d'achats/ventes nécessaires) de façon à respecter les proportions initiales retenues par l'allocation stratégique. Cette approche est également appelée règle de gestion par « rebalancement » ou *Fixed-mix*. Il s'agit d'une stratégie d'investissement dont le principe consiste dans un rebalancement périodique des poids vers des niveaux cibles, correspondants à ceux de l'allocation stratégique optimale.

Cette stratégie repose implicitement sur le principe suivant : à l'issue de chaque sous-période et lors du rebalancement, la partie des actifs qui a été la plus performante durant cette sous-période est vendue suite à l'augmentation de son poids par rapport aux niveaux cibles de l'allocation stratégique. Elle est donc vendue à un prix considéré comme « élevé ». L'autre partie des actifs va être, quant à elle, achetée à un prix « pas cher » : son rendement étant relativement plus faible sur la sous-période considérée. Nous convergeons finalement vers une composition du portefeuille identique à celle de l'allocation stratégique cible (cf. Zenios [2007]). Cela constitue un point fort de cette stratégie. Les propriétés théoriques de la stratégie *Fixed-Mix* sont par exemple discutées dans Merton [1990], Dempster et al. [2003], Infanger [2002], Brinson et al. [1991] et Kouwenberg [2001].

Il s'agit d'une stratégie d'une grande importance en pratique puisqu'elle est souvent utilisée pour le choix d'un portefeuille de référence appelé *benchmark* : ce dernier permet de juger la performance des gérants. Plusieurs études ont été dédiées à l'évaluation de la contribution de la gestion active des actifs (arbitrages sur de courtes périodes) dans la performance globale du portefeuille (cf. Brinson et al. [1991]). Les résultats empiriques montrent la faible valeur ajoutée de la gestion active et confirment l'importance du benchmark (et donc l'importance de la stratégie *Fixed-Mix* qui définit ce dernier) dans le processus de gestion (cf. Blake et al. [1999], Brinson et al. [1991]).

La combinaison de la stratégie *Fixed-Mix* avec un modèle d'ALM stochastique dédié à un régime de retraite typique en France sera l'objet de l'étude suivante. Le GSE utilisé est celui décrit dans la partie I de ce travail.

I-2 Application

Lors de cette application nous présentons une modélisation du régime-type. Deux études sont également menées : une première étude sur des critères d'optimisation pouvant être considérés et une deuxième étude sur la question du taux d'actualisation des engagements d'un régime de retraite. Ensuite, nous présentons les hypothèses et la méthodologie retenues. Les résultats obtenus sont enfin discutés.

I-2-1 Modélisation du régime-type

Cette sous-section a pour objet de présenter une proposition de modélisation des flux financiers futurs dans le contexte de la gestion à long terme d'une réserve.

I-2-1-1 *Notations*

Dans la suite, nous utiliserons les notations suivantes :

n le semestre ;

$(\rho_n)_{n\geq 0}$ la suite des rendements (nets de frais) de l'actif du régime où ρ_n désigne le rendement (*a priori* aléatoire) sur la période $[n, n+1[$;

$(p_n)_{n\geq 0}$ la suite des prestations désinflatées où p_n représente le montant des prestations à verser au titre de la période (semestre) $[n, n+1[$;

$(c_n)_{n\geq 0}$ la suite des cotisations désinflatées où c_n représente le montant des cotisations à recevoir au titre de la période $[n, n+1[$;

$\left(i_n^{(x)}\right)_{n\geq 0}$ la suite des facteurs d'indexation de la grandeur x désinflatée avec $x = c, p$ selon que nous considérons les facteurs d'indexation des cotisations ou des prestations ;

$\nabla^{(x)}$ l'indicateur de paiement de la suite associée à la grandeur x : $\nabla^{(c)} = 0$ si les cotisations relatives à la période $[n, n+1[$ sont versées en n, $\nabla^{(c)} = 1$ si elles sont versées en $n+1$.

I-2-1-2 *Modélisation*

L'objet de ce paragraphe est de préciser les grandeurs qui devront être suivies afin de construire les indicateurs / critères permettant de réaliser l'allocation stratégique d'actifs. À ce stade de l'étude, il ne s'agit toutefois pas de préciser quels pourront être ces indicateurs et ces critères.

i. Dynamique du régime

Du point de vue comptable, le montant de la réserve totale évolue selon la dynamique suivante :

$$R_{n+1} = \left[R_n + c_{n-\nabla^{(c)}} \times i^{(c)}_{n-\nabla^{(c)}} - p_{n-\nabla^{(p)}} \times i^{(p)}_{n-\nabla^{(p)}} \right] \times (1+\rho_n) + c_{n-\nabla^{(c)}+1} \times i^{(c)}_{n-\nabla^{(c)}+1} - p_{n-\nabla^{(p)}+1} \times i^{(p)}_{n-\nabla^{(p)}+1}$$

Voici un exemple de format d'alimentation des données de la maquette d'ALM :

Période	Flux de cotisations	Actif Circulant	1	2	3	4	5
1	$\overline{c_1}$		$\overline{p_1}$	D_1^2	D_1^3	D_1^4	D_1^5
2	$\overline{c_2}$		-	$\overline{p_2}$	D_2^3	D_2^4	D_2^5
3	$\overline{c_3}$		-	-	$\overline{p_3}$	D_3^4	D_3^5
4	$\overline{c_4}$			-	-	$\overline{p_4}$	D_4^5
5	$\overline{c_5}$		-	-	-	-	$\overline{p_5}$

Tab. 28 : *Illustration du modèle d'alimentation des données de la maquette d'ALM*

- $\overline{c_i}$: cotisation totale reçue à la date *i*.
- $\overline{p_i}$: prestation totale payée à la date *i*.
- D_j^i : droit acquis exigible à la période *j* au titre des cotisations passées cumulées jusqu'à la période *i* (autrement dit les stocks à *i* des droits acquis et exigibles en *j*)

ii. Suivi comptable

Le suivi comptable doit permettre de reconstituer, à chaque début et fin de période, le bilan du régime. Celui-ci est composé :

- à l'actif :
  - de la valeur comptable des actifs financiers ;
  - d'une éventuelle créance (nette) envers des débiteurs au titre des cotisations à percevoir ;
  - des autres actifs non-techniques ;
  - de la trésorerie ;
- au passif :
  - de la valeur comptable de la réserve ;
  - des provisions techniques complémentaires ;
  - d'une éventuelle dette (nette) envers ;
  - des autres passifs non-techniques (ex transfert, frais de gestion administratives futures etc.).

Par ailleurs, le rendement comptable de l'actif du régime étant constitué :

- des variations de surcote-décote des titres obligataires,
- des tombées de coupons des titres obligataires,

- des tombées de dividendes des actions et des OPCVM,

- des loyers des investissements immobiliers,

- de la rémunération de la trésorerie,

- de la réalisation des plus/moins-values latentes.

iii. Suivi de la trésorerie

Sur chaque période, le besoin en trésorerie est obtenu comme étant le solde entre :

- d'une part les prestations payées sur la période et les frais sur la période ;

- d'autre part les cotisations perçues sur la période et les revenus des actifs financiers (y compris intérêts provenant d'éventuelles créances sur l'État et les collectivités) avant tout désinvestissement / investissement.

S'il est positif, le besoin en trésorerie permet de définir le niveau minimal de réalisations d'actifs à effectuer pour être capable de payer les prestations.

iiii. Suivi de la valeur

Il s'agit ici d'effectuer, à chaque début et fin de période, un bilan « économique » du régime en valorisant chacun de ses postes, non plus selon les méthodes comptables qui s'y appliquent mais selon des méthodes économiques.

En pratique, les principaux postes impactés sont :

- les actifs financiers qui sont valorisés à leur prix de marché ou selon une méthode *market-consistent* pour les actifs non cotés sur un marché liquide et actif ;

- l'engagement du régime au titre des droits acquis évalué selon une approche Solvabilité II (*best estimate* + marge pour risque) ;

- l'engagement du régime au titre des droits futurs (si les conditions futures du régime sont déséquilibrées).

Ces deux derniers points ne pourront toutefois être appréciés que selon la disponibilité des données sur la distinction entre les droits acquis et les droits futurs. Nous supposerons que ces informations ne sont pas disponibles à ce stade. Dans le cadre de la définition des inputs des modèles développés, il conviendra toutefois de distinguer la population en groupe fermé (en y distinguant les rentiers actuels et les rentiers futurs) et les nouveaux entrants.

Pour l'établissement des principes de cette comptabilité « économique », la référence aux modèles IFRS et Solvabilité II pourra être utilisée pour justifier des méthodes et définir un cadre.

I-2-2 Etude des critères d'allocation

Nous visons, dans ce qui suit, à formaliser et à synthétiser certains critères (qui peuvent servir en tant que contraintes ou en tant qu'objectifs (c'est-à-dire la fonction à optimiser) pour un

régime de retraite. Pour mémoire, l'allocation stratégique doit permettre sur différents horizons d'optimiser la fonction objectif, en respectant l'ensemble des contraintes et en tenant compte des règles de gestion. Une fonction « objectif » est ainsi à définir. Elle reflètera le « budget de risque » de l'investisseur. La forme générale de cette fonction sera la maximisation de la valeur d'un indicateur de performance sous contrainte que la valeur d'un indicateur de risque soit inférieure à un plafond. La valeur généralement utilisée dans un cadre d'étude d'ALM financière est celle du ratio de financement : il s'agit du rapport à un instant donné entre, d'une part, la valeur de la réserve ajustée à la marge d'exploitation et d'autre part, la valeur actualisée des engagements futurs (cf. Planchet et al. [2007]).

I-2-2-1 *Analyse générale de la solvabilité du régime*

- **Approche code des assurances**

L'analogie avec le code des assurances et, en particulier, des régimes de rentes provisionnés, conduit à s'assurer que :

$$R_n \geq PM_n = \mathbf{E}_n\left(\sum_{k=n}^{+\infty}\left[p_{k-\nabla^{(p)}} \times i^{(p)}_{k-\nabla^{(p)}}\left(1+r_a\right)^{-k+n} - c_{k-\nabla^{(c)}} i^{(c)}_{k-\nabla^{(c)}}\left(1+r_a\right)^{-k+n}\right]\right)$$

Où :

r_a représente le facteur d'actualisation exprimé dans la base de la période retenue ;
n représente la date d'observation ;
$\mathbf{E}_n$ représente l'espérance à la date n.

Ce test permet de s'assurer que la valeur actuelle des soldes techniques futurs est couverte par le montant de la réserve. Ce qui signifie que, sous l'hypothèse que les flux futurs anticipés se réaliseront, placer le montant de la réserve au taux d'actualisation r_a permettra de régler les prestations dues. Incidemment ce critère peut être réalisé sans que pour autant le montant de la réserve soit positif à tout instant dans le futur : à certains moments, le solde peut être débiteur et nécessiter un emprunt de trésorerie.

Une version étendue de ce critère consiste à regarder ce solde à chaque date de fin de période m : $\forall m \geq n, R_m \geq PM_m = \mathbf{E}_n\left(\sum_{k=m}^{+\infty}\left[p_{k-\nabla^{(p)}} \times i^{(p)}_{k-\nabla^{(p)}}\left(1+r_a\right)^{-k+n} - c_{k-\nabla_c} \times i^{(c)}_{k-\nabla_c}\left(1+r_a\right)^{-k+n}\right]\right)$

- **Approche économique**

Sur la base du suivi de la « valeur » du régime (cf. la sous-section I-2-1 de ce chapitre), il apparaît opportun d'estimer le niveau du surplus selon une approche économique et d'en suivre l'évolution au fil du temps.

Dans ce contexte, le surplus économique s'exprime comme étant la différence entre l'actif du régime (évalué économiquement ici) et les provisions du régime (également évaluées selon une approche économique) et autres dettes.

- **Objectifs de l'allocation**

L'allocation stratégique repose souvent sur la maximisation de la fonction de la richesse du régime (valeur de la réserve) ou de son corollaire le Taux de Rendement Interne (TRI), particulièrement sur différents horizons d'investissement (court terme, moyen terme, long terme). Le TRI est le taux de rendement permettant d'avoir la relation suivante :

$$R_{t0} = \sum_{t=t0+1}^{d_H-1} \frac{\overline{c_i} - \overline{p_i}}{(1+TRI)^{i-t0}} + \frac{R_{d_H}}{(1+TRI)^{d_H-t_0}}$$

où :

$\overline{c_t}$ représente les cotisations (y compris inflation) à la date *t* ;

$\overline{p_t}$ représente les prestations (y compris inflation) à la date *t* ;

R_{t_0} représente la valeur de la réserve à la fin de l'année t_0, date de détermination de l'allocation stratégique (date de départ) telle que définie ci-dessus;

d_H représente l'horizon de placement de l'allocation stratégique ;

TRI le taux de rendement interne.

ce qui revient à :

$$R_{t0}(1+TRI)^{d_H-1} - \sum_{i=t0+1}^{d_H-1}(\overline{c_i} - \overline{p_i})(1+TRI)^{d_H-i} = R_{d_H}$$

Il est précisé que la réserve utilisée ici est la réserve investie (réserve financière).

Dans la description des critères recherchés de l'allocation stratégique, il convient de distinguer différents types de régimes de retraite (régime provisionné, régime partiellement provisionné, régime non provisionné). En outre, quel que soit le régime, dans le développement de l'allocation stratégique de l'actif, une attention particulière sera portée à la distinction entre la phase de constitution et la phase de restitution du régime.

I-2-2-2 *Critères d'allocation en fonction du type du régime*

Dans ce qui suit, nous présentons certains critères qui peuvent servir à la fois comme objectif et/ ou contrainte selon le type du régime.

- **Pour un régime partiellement provisionné :** maintenir un ratio de couverture des droits acquis par la réserve. Il s'agit du rapport entre la valeur des actifs de la réserve financière et la valeur des flux futurs (au titre des droits acquis) actualisés au taux sans risque. Les contraintes de solvabilité pour l'allocation stratégique peuvent être de différentes ordres : elles portent d'une part sur l'horizon de viabilité, et d'autre part sur le niveau des réserves et des prestations.

Un horizon de viabilité minimal de H_v *années*

L'horizon de viabilité correspond à l'horizon d'épuisement des réserves (pour mémoire, la réserve à la date *t* est égale à la somme de la réserve à la date *t*-1, de la marge d'exploitation à

la date *t* et des produits financiers à la date *t*). Ce critère revient à vérifier la condition suivante (pour H_v=30):

$$\frac{R_t + \sum_{i=t+1}^{t+30} \frac{\overline{c_i}}{\prod_{j=t+1}^{i}(1+r_j)}}{\sum_{i=t+1}^{t+30} \frac{\overline{p_i}}{\prod_{j=t+1}^{i}(1+r_j)}} \geq 1.$$

Un horizon de solvabilité minimal de h années

Un deuxième critère possible d'allocation stratégique porte sur le rapport entre le niveau des prestations à terme à h ans et les réserves à cette échéance : la réserve devra assurer, de façon permanente, la capacité de provisionner *m* fois la prestation à payer dans *h* années. Autrement dit, pour chaque année *t*, la somme de la réserve de l'exercice ajustée de l'actualisation des flux futurs du régime, projetées sur les *h* années suivantes doit au moins être équivalente à un engagement prédéfini (par exemple une fois et demi l'actualisation en *t* de la prestation prévisionnelle à horizon de h ans par rapport à *t*).

Le taux d'actualisation r_t utilisé doit correspondre à un taux de rendement prudemment estimé des actifs en représentation de l'engagement sur la période de projection. Les taux retenus peuvent ainsi être ceux de la courbe des taux *forwards* zéro-coupon obtenue à partir des obligations état indexées.

Financièrement ce type de critère peut être exprimé pour chaque fin d'année par la formule suivante :

$$R_t + \sum_{i=t+1}^{t+h} \frac{\overline{c_i} - \overline{p_i}}{\prod_{j=t+1}^{i}(1+r_j)} \geq m \times \frac{\overline{p_{t+h}}}{\prod_{j=t+1}^{t+h}(1+r_j)}$$

où :

- $\overline{c_t}$: représente les cotisations (y compris inflation) à la date *t* ;
- $\overline{p_t}$: représente les prestations (y compris inflation) à la date *t* ;
- r_j : représente le taux de rémunération prudent de la réserve à la date j (par convention nous l'assimilons au taux de rendement de l'obligation indexée de long terme) ;
- m : représente le rapport minimum entre la valeur de la réserve à une date donnée et le flux de prestation cible actualisé correspondant (par exemple m = 1,5) ;
- h : représente l'horizon de projection pour le flux de prestation à couvrir par rapport à la date t.

Ce critère peut être également exprimé par la relation suivante :

$$\left[C \geq m\right], \text{ avec } C = \left(\frac{R_t + \sum_{i=t+1}^{t+h} \frac{\overline{c_i} - \overline{p_i}}{\prod_{j=t+1}^{i} \left(1+r_j\right)}}{\frac{\overline{p_{t+h}}}{\prod_{j=t+1}^{t+h} \left(1+r_j\right)}} \right).$$

Dans un univers incertain (en particulier avec un horizon de long terme), le respect du critère doit être probabilisé avec un niveau de confiance $1-\alpha$ (par exemple 97,5 %) où α est le risque d'erreur toléré ou « la probabilité de ruine », soit :

$$P\left[C \geq m\right] \geq 1-\alpha$$

Ainsi, pour les critères de solvabilité associés à la comparaison entre le niveau des réserves et celui des prestations l'exemple de formule développée ci-dessus donne :

$$P\left[\left(\frac{R_t + \sum_{i=t+1}^{t+h-1} \frac{\overline{c_i} - \overline{p_i}}{\prod_{j=t+1}^{i} \left(1+r_j\right)}}{\frac{p_{t+h}}{\prod_{j=t+1}^{t+h} \left(1+r_j\right)}} \right) \geq m \right] \geq 1-\alpha$$

- **Pour un régime provisionné** : maintenir un ratio de couverture des droits acquis par le total de la réserve et de la marge actuarielle future (ou « Potentiel de Répartition Futur », cf. Delarue [2001]). Il s'agit du rapport entre la valeur des actifs de la réserve financière et la valeur des flux futurs (au titre des droits acquis et des cotisations futures nettes des prestations induites par les droits que ces cotisations génèrent) actualisés au taux sans risque. Ce ratio peut-être fixé entre 90 % et 100 % par exemple.

Autrement dit, dans le cas d'un régime dont le but est de provisionner les engagements futurs, nous cherchons à rémunérer au mieux ces droits acquis en maximisant le rendement des réserves représentatives des droits acquis par les cotisants du régime.

L'idée est donc de s'intéresser d'abord pendant la durée de l'allocation à l'engagement économique du régime, c'est-à-dire aux prestations futures liées aux droits acquis en vigueur et évaluées à partir du rendement technique du régime. L'allocation stratégique choisie sur la durée *D* doit permettre de minimiser le coût de financement de ces engagements sur la période considérée.

Aussi, le critère proposé consistera à minimiser l'espérance de la valeur actuelle des Droits Acquis à la date t_0. Cette valeur actuelle est la somme :

- des prestations dues et réglées sur la période d'allocation (t_0 à *D*) actualisées aux taux de rendement périodiques du portefeuille financier (allocation d'actifs) ;

- de la provision de tous les flux de prestations dues au-delà de cet horizon représentatifs des droits acquis jusque là et actualisés au taux sans risque.

où :

$$ValeurDroitsAcquis_{t0} = \sum_{i=t0+1}^{d_H-1} \frac{\sum_{j=t0+1}^{i} D_j^i}{\prod_{j=t0+1}^{i}\left(1+r_j\right)} + \frac{Provision_TSR_{31/12/d_H}}{\prod_{j=t0+1}^{d_H-t0}\left(1+r_j\right)}$$

$$Provision_TSR_{31/12/d_H} = \sum_{i=t0+1}^{d_H-1} \frac{D_{d_H}^i}{\prod_{i=d_H}^{d_f}\left(1+r_{sr}\right)}$$

t_0 : date initiale (par exemple date de calcul de l'allocation stratégique) ;

D_j^i : représente les prestations (avec l'indexation inflation) payables à la date i correspondants aux stocks en j des droits acquis et non échus antérieurs à cette date ;

r_j : représente le rendement du portefeuille d'actifs financiers au cours de la période j ;

d_H : représente l'horizon de l'allocation stratégique ;

$Provision_TSR_{31/12/d_H}$: représente la valeur des flux futurs (au titre des droits acquis) actualisés au 31/12/ d_H au taux sans risque.

d_f : date finale de la série disponible des flux au titre des droits acquis à t_0 (nous choisirons par exemple une date lointaine, ou bien un horizon de « viager », i.e la date de décès du dernier survivant des cotisants).

Concernant les contraintes relatives aux taux de couverture d'autres engagements tels que les droits acquis. Les contraintes à respecter seront par exemple :

- couverture des prestations acquises par la réserve :

$$\varphi_t = \frac{Réserves\left(31/12/t\right)}{Provisions_TSR\left(31/12/t\right)} > \beta$$

avec :

φ_t est le ratio de couverture à la date t ;

β est compris entre 6 % et 10 % (exemple régime cas 1) ;

$Provisions_TSR(31/12/t)$ représente la valeur des flux futurs (au titre des droits acquis) actualisés au 31/12/t au taux sans risque.

- couverture des prestations acquises par la réserve et la marge actuarielle future (ou « Potentiel de Répartition Futur ») : Si nous définissons la Marge Actuarielle

Future (MAF) par : $MAF = \sum_{i=t}^{\infty} \frac{\bar{c}_i - \bar{p}_i}{\prod_{j=t}^{i}(1+r_j)}$, le ratio de couverture des prestations peut s'exprimer par :

$$\phi_t = \frac{Réserves(31/12/t) + MAF}{Provisions_TSR(31/12/t)}.$$

En résumé, les contraintes à considérer pour les régimes de type provisionnés avec prise en compte des droits acquis sont les suivantes (au choix) :

- la *VaR* à 99 % du ratio de couverture des prestations acquises par la réserve doit être supérieure à un coefficient béta1 (fixé en input, exemple entre 6 % et 10 %)

VaR99% [*Réserves*(*31/12/t*) / *Provisions_TSR*(*31/12/t*)] > *béta 1*

- la *VaR* à 99 % du ratio de couverture des prestations acquises par la réserve et la marge actuarielle future doit être supérieure à un coefficient béta 2 (fixé en input, exemple entre 90 % et 100 %)

VaR99% [(*Réserves*(*31/12/t*) + *MAF*) / *Provisions_TSR*(*31/12/t*)] > *béta 2*

o **Pour un régime non provisionné** : un principal objectif dans ce cas vise à maximiser un ratio de couverture des prochains flux nets de règlement par les actifs jusqu'à la date prévisionnelle d'épuisement des réserves, supérieur à 100 %. Il s'agit du rapport entre le total de la valeur des actifs de la réserve financière augmenté de la valeur actualisée des flux futurs de règlement nets encaissés (cotisations moins prestations) et la valeur des flux futurs de règlement nets décaissés. Il s'agit en fait d'un ratio de « liquidité » ou de « trésorerie » du régime.

Les contraintes posées consistent, par exemple, à vérifier que dans la majorité des cas (par exemple dans 99 % des cas) : le ratio de couverture des flux futurs de trésorerie ou règlement (amputé d'une année par rapport à l'horizon de viabilité du régime qui correspond à l'horizon d'épuisement des réserves) est supérieur à 100 %, cela signifie que le budget de risque pour le choix de l'allocation ne peut être supérieur à la perte d'une année de prestation.

Dans ce cadre, nous supposons que la Valeur à Risque (VaR), pour un certain niveau de confiance, du rapport (ratio de couverture) à la date d_H entre les deux quantités ci-après doit être supérieure à 100 % :

- la réserve, augmentée de la valeur actualisée des flux futurs de règlement net encaissés (i.e. positifs) jusqu'à la date $d_e - 1$
- la valeur actualisée des flux futurs de règlement net (i.e. négatifs) jusqu'à la date $d_e - 1$

Le taux d'actualisation des flux futurs est le même que celui défini pour l'actualisation des flux d'un régime partiellement provisionné lors du calcul des ratios prudentiels.

I-2-2-3 *Critères d'allocation en fonction da la phase dans laquelle évolue le régime*

Dans le cas où le régime est dans une phase d'accumulation de réserve l'horizon naturel pourra être la fin prévisionnelle de la période d'accumulation des réserves. Par ailleurs, pour un régime en phase d'épuisement, la maximisation du ratio de liquidité peut rester un bon objectif. Dans le cas d'un régime de ce type, la fonction objectif à maximiser pourra ainsi être formalisée de la manière suivante :

Médiane du rapport à la date d_H *entre les deux quantités ci-après :*

- *le total de la valeur des actifs de la réserve financière, augmentée de la valeur actualisée des flux futurs de règlement net encaissés (i.e. positifs) jusqu'à la date* d_e
- *la valeur actualisée des flux futurs de règlement nets décaissés (i.e. négatifs) jusqu'à la date* d_e

avec :

d_H = date de l'horizon de l'allocation

d_e = date prévisionnelle d'épuisement des réserves

Nous rappelons que le flux de règlement net pour une année *i* est la différence entre les cotisations encaissées de l'année *i* et les prestations versées de l'année *i* soit $\overline{c_i} - \overline{p_i}$.

Quel que soit le type du régime ou la phase dans laquelle il évolue, une attention particulière doit être apportée aux pertes potentielles de court terme.

I-2-3 Etude de la question du taux d'actualisation des engagements

Deux principaux types d'études de congruence actifs/passifs (ALM) peuvent être présentés. Leur différence clé se situe au niveau du choix du **taux d'actualisation des engagements**.

ALM actuarielle	*ALM financière (market-based)*
• L'évaluation des engagements se fait en appliquant un taux d'intérêt fixe (appelé taux technique ou taux actuariel).	• L'évaluation des engagements se fait en utilisant les taux du marché (taux sans risque, taux de swap).
• L'ALM actuarielle permet d'obtenir une valeur « actuarielle » des engagements	• L'ALM financière permet d'obtenir une valeur de « marché » des engagements.
• Dans ce cadre, les engagements ne contiennent pas de risque de taux.	• Dans ce cadre, le profil de risque des engagements ressemble à un portefeuille obligataire.
→ Notre benchmark sera le « rendement absolu »	→ Benchmark: « engagements »

Les conséquences de ces études sur la stratégie d'investissement optimale peuvent être synthétisées comme suit :

ALM actuarielle	*ALM financière (market-based)*
la stratégie optimale cherche à atteindre un rendement égal ou supérieur au rendement actuariel requis avec la volatilité absolue la plus faible possible.	la stratégie optimale se compose de deux portefeuilles: – Un portefeuille **core** «répliquant» les engagements au plus près (portefeuille d'obligations de longues durations) : recherche de couverture des engagements – Un portefeuille **satellite** exposant les actifs à des primes de risque supérieures à celles des obligations (actions, immobilier, actifs alternatifs) : recherche de performance absolue

Le choix du taux d'actualisation des engagements suscite un débat entre les professionnels depuis quelques années, surtout depuis la crise des fonds de pension de 2001 (crise de sous financement). Un écart de quelques points dans le choix du taux peut faire varier le montant de la provision de manière significative. Certains recommandent aussi d'utiliser des taux de court terme pour les provisions d'une population proche de l'âge de la retraite, et des taux de long terme pour une population jeune. Dans tous les cas, le débat sur le taux d'actualisation ne peut pas être décorrélé du processus de réévaluation des rentes (cf. Planchet et Thérond [2007]).

Dans le cas où l'actualisation est effectuée avec le taux sans risque du marché, le risque de taux intérêt se manifeste surtout en cas de baisse des taux. Quand le taux d'intérêt baisse, la valeur de l'actif augmente, dans des proportions qui dépendent de la composition du portefeuille du fonds (actions ou obligations notamment). Si ce taux d'intérêt est celui utilisé pour actualiser les engagements, leur valeur actualisée augmente également.

Le risque de taux provient de ce que l'actif et le passif n'ont généralement pas la même sensibilité aux taux d'intérêt : une baisse des taux à long terme se traduit, en règle générale, par une hausse de la valeur du passif plus importante que celle de l'actif. Dans ces conditions, le taux de couverture du fonds, défini comme le rapport entre la valeur des actifs et la valeur actualisée des engagements, se détériore. De même, la baisse des taux réduit la marge financière future car le rendement des nouveaux actifs peut apparaître insuffisant pour faire face aux taux garantis par les contrats antérieurs.

Jusqu'à une date récente, les engagements des fonds de pension anglo-saxonnes (Etats-Unis, Royaume-Uni, Pays-bas, etc.) étaient valorisés avec un taux d'actualisation fixe, par exemple 4 % par an. Mais, depuis quelques années, les règles prudentielles et les normes comptables de ces pays ont évolué en direction d'une comptabilisation en juste valeur où la courbe des taux d'intérêt de marché est utilisée pour actualiser les versements (selon les pays, il s'agit du taux des obligations d'État, du taux moyen des obligations émises par les entreprises ou du taux observé sur le marché des contrats d'échange de taux appelé taux de *swap*). La mesure en juste valeur présente simultanément des avantages et des inconvénients :

- **Avantages** de la mesure en juste valeur des engagements des fonds de pension :
Elle accroît la transparence, favorise la portabilité des droits à la retraite, qui peuvent être transférés à tout moment d'un fonds à l'autre sur la base de leur valeur de marché, et permet d'utiliser des instruments financiers pour gérer plus finement les risques.

- **Inconvénients** de la mesure en juste valeur des engagements des fonds de pension :
Le passage d'un taux fixe à des taux de marché crée un risque de taux sur le passif du fonds, qui n'est généralement pas couvert par une variation équivalente de la valeur de l'actif.

D'autre part, la hausse des taux réduit la valeur de marché des actifs, en particulier celle des portefeuilles obligataires, et peut provoquer éventuellement le rachat de certains contrats, notamment ceux à rendement garanti inférieur à celui des nouveaux contrats proposés.

Dans ce qui suit, nous allons illustrer la mise en place d'un modèle d'allocation stratégique d'actifs pour un régime de retraire partiellement provisionné. Ce modèle est basé sur la stratégie *Fixed-Mix*.

I-2-4 Hypothèses du modèle

Concernant le passif, le fonds est considéré comme un fonds fermé (pas de nouveaux cotisants jusqu'à la fin de la période d'étude). Les flux nets à payer sont définis en termes de « marge d'exploitation » :

+ Flux techniques (cotisations normales et rétroactives + compensations)
- Prestations (allocations + capitaux décès + transferts)
= **Marge technique**
- Frais de gestion
- Dépenses au titre de l'action sociale
= **Marge d'exploitation**

Nous supposons que la projection du flux est déjà effectuée en termes de flux réels. La période de projection sera celle du 01/01/2010 au 31/12/2060.

Le graphique 41 montre un exemple de projection des flux de passif pour ce type de fonds. L'horizon de projection des scénarios économiques et financiers est de 50 années (du 01/01/2010 au 01/01/2060) avec un pas de projection trimestriel. La fréquence de rebalancement du portefeuille vers l'allocation stratégique testée est également trimestrielle (nous rappelons qu'on se situe dans le cas d'un modèle *Fixed-Mix* et ainsi un retour périodique aux poids prédéfinis par l'allocation stratégique est effectué).

Une approche d'ALM financière a été adoptée : actualisation des flux du passif avec la courbe des taux simulée à chaque instant. La valeur actuelle à l'instant t de N flux futurs est donnée par l'expression suivante :

$$Engagements\ Actualisés_t = \sum_{i=1}^{N} \frac{Flux_i}{\prod_{j=1}^{i}(1+r_j)}$$

avec : r_j le taux d'intérêt sur la sous période j (taux réel en cas de flux réels et taux nominal en cas de flux nominaux).

La même démarche est utilisée pour l'indexation des flux de passif par les taux d'inflation simulés. Dans ce cas, le flux réel à un instant *t* est multiplié par le coefficient d'indexation déduit à partir des taux d'inflation réalisés sur les périodes entre l'instant initial t_0 =0 et *t* :

$$Flux\ indexé_t = Flux\ Réel_t \times \prod_{j=1}^{t} \left(1 + i_j\right)$$

avec : i_j le taux d'inflation sur la sous période *j*.

Nous supposons que le choix de l'allocation stratégique sera principalement guidé par des critères portant sur le niveau des réserves : maximisation d'un montant de réserve à terme ou de sa résultante le Taux de Rendement Interne (TRI) à horizon de l'allocation stratégique.

Autrement dit, nous supposerons que la politique d'investissement vise à maximiser la valeur de la réserve financière du régime à une date donnée dans le futur : ce qui revient à maximiser le taux de rendement interne de la réserve entre la date initiale et cette date future.

Simultanément, deux contraintes seront à respecter par l'allocation optimale :

Contrainte 1 : Le rapport à une date t_0 donnée entre :

- La somme de la valeur de la réserve à la date t_0 et des flux futurs de cotisation sur les H_v prochaines années (ces flux sont *actualisés* à la date t_0)
- La somme des flux futurs de prestation sur les H_v prochaines années (ces flux sont *actualisés* à la date t_0) devra également être supérieur à un niveau m_1 dans $(1-\alpha_1)$ % des cas.

Contrainte 2 : Le rapport à une date t_0 donnée entre :

- La valeur projetée de la réserve à la date t_0 plus H_{max} années
- La prestation prévisionnelle versée à la date t_0 plus H_{max} années devra être supérieur à un niveau m_2 dans $(1-\alpha_2)$ % des cas

De point de vue mathématique, ce système d'optimisation peut être formulé ainsi :

$$\underset{TRI}{Max}\, R_{t0}(1+TRI)^{H_{\max}} + \sum_{i=t0+1}^{(t_0+H_{\max})-1} (\overline{c_i} - \overline{p_i})(1+TRI)^{(t_0+H_{\max})-i} = R_{(t_0+H_{\max})}$$

$$\text{S/C} \qquad P\left[\frac{R_t + \sum_{i=t_0+1}^{t_0+H_v} \frac{\overline{c_i}}{\prod_{j=t_0+1}^{i}(1+r_j)}}{\sum_{i=t_0+1}^{t_0+H_v} \frac{\overline{p_i}}{\prod_{j=t_0+1}^{i}(1+r_j)}} \geq m_1\right] \geq 1-\alpha_1$$

$$P\left[\left(\frac{R_t + \sum_{i=t_0+1}^{t_0+H_{\max}} \frac{\overline{c_i} - \overline{p_i}}{\prod_{j=t_0+1}^{i}(1+r_j)}}{\frac{\overline{p_{t_0+H_{\max}}}}{\prod_{j=t_0+1}^{t_0+H_{\max}}(1+r_j)}}\right) \geq m_2\right] \geq 1-\alpha_2$$

Avec :

$\overline{c_t}$ représente les cotisations (y compris inflation) à la date *t* ;

$\overline{p_t}$ représente les prestations (y compris inflation) à la date *t* ;

R_{t_0} représente la valeur de la réserve à la fin de l'année t_0, date de détermination de l'allocation stratégique (date de départ)

TRI le taux de rendement interne.

r_j : représente le taux de rémunération prudent de la réserve à la date j (taux nominal sans risque);

m_1 : représente le rapport minimum pour la contrainte 1 (dans notre cas $m_1 = 1$) ;

m_2 : représente le rapport minimum pour la contrainte 2 (dans notre cas $m_2 = 1{,}5$) ;

H_v : représente l'horizon considéré pour le calcul du rapport de la contrainte 1 (dans notre cas H_v =30 ans).

H_{max} : représente l'horizon considéré pour le calcul du rapport de la contrainte 2 (dans notre cas H_{max} =20 ans).

α_1 et α_2 correspondent aux marges d'erreur au niveau du respect de la contrainte 1 et 2 respectivement ($\alpha_1 = \alpha_2$ =2.5 % dans notre cas)

Le rapport de la première contrainte peut être assimilé à un ratio de couverture des flux de prestation sur les prochaines années par la réserve initiale augmentée des flux futurs de cotisation sur les mêmes années : cela assure dans un certain sens la **viabilité** du régime. De même, le rapport de la deuxième contrainte représente un ratio de couverture de la prestation à payer dans années par la valeur de la réserve de la même année : cela assure dans un certain sens un niveau de **solvabilité** requis par le fonds après H_{max} années. Cela correspond au cas d'un régime de retraite **partiellement provisionné**.

Concernant les contraintes de poids sur la composition de l'allocation cible, Il doit être possible d'imposer *a priori* pour chaque classe d'actif de l'allocation une composition minimum et maximum à respecter pour chaque allocation testée.

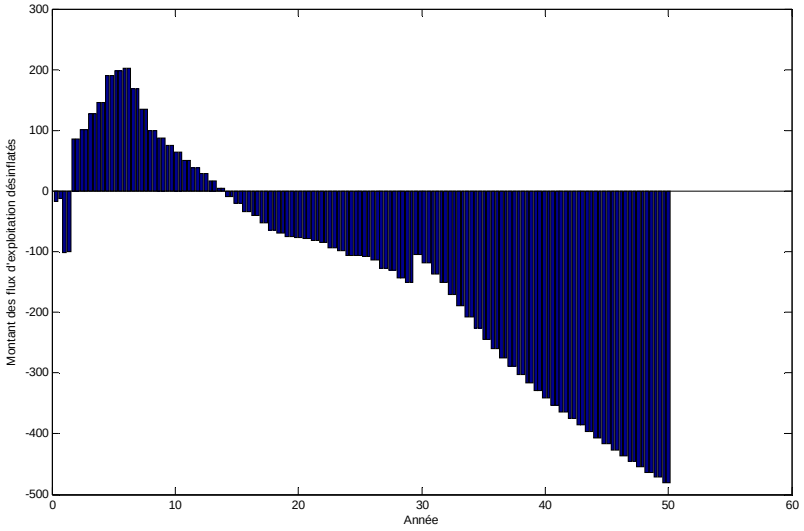


Fig. 41 : *Exemple d'évolution des flux nets de trésorerie (flux désinflatés)*

Nous rappelons finalement, qu'en termes de règles de gestion, il est considéré qu'à chaque fin de période de simulation nous « forçons » le fonds à respecter les proportions initiales (« rebalancement »). Ceci est du à la stratégie *Fixed-Mix* adoptée. Cette dernière présente l'avantage :

- d'être cohérente avec l'idée d'allocation stratégique ;
- de considérer le fonds comme un FCP (Fonds Commun de Placement) en travaillant avec son rendement issu des pondérations des rendements de chaque poche ;
- d'éviter de définir des règles complexes d'investissement / désinvestissement.

I-2-5 Méthodologie retenue

Les étapes de cette étude actif-passif sont les suivantes :

Etape de simulation des scénarios économiques et financiers

Il s'agit d'une étape commune et préliminaire et au cours de laquelle nous nous basons sur le GSE proposé dans la partie I pour :

- La simulation des différents taux d'intérêt réels (longs et courts), des taux d'inflation anticipés (longs et courts) avec déduction des taux nominaux (longs et courts).
- La simulation du rendement des actions, du rendement de l'immobilier et de l'évolution du taux monétaire.
- La détermination des prix des zéros coupons nominaux, indexés à l'inflation et à risque de crédit.
- La détermination des performances des produits de taux, selon les formules présentées dans l'annexe relatif aux scénarios déterministes.

Ces simulations se feront selon la structure schématique linéaire (quel que soit la classe d'actifs considérée). Ce type de structure est souvent utilisé dans le cas de la stratégie *Fixed-Mix* (cf. Kouwenberg [2001]). Pour rappel, cette structure se présente ainsi (cas des actions) :

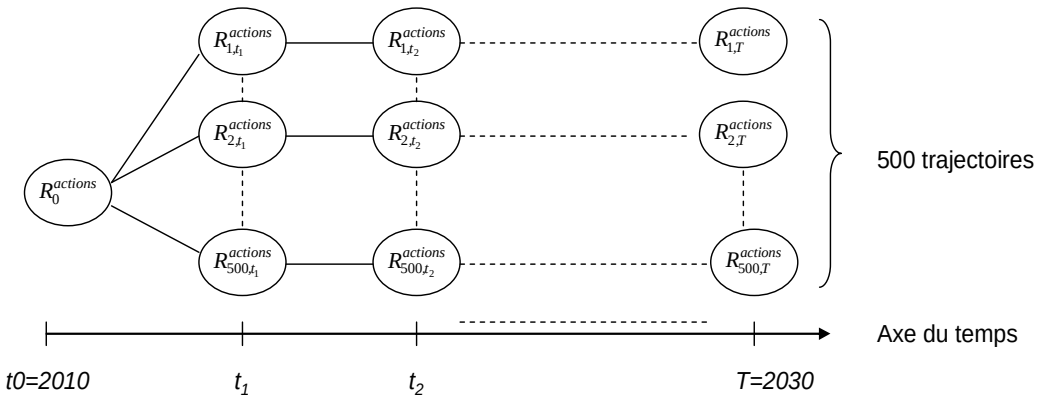


Fig. 42 : *Illustration de la structure schématique linéaire dans le cas de la projection des rendements des actions*

<u>Etape de préparation des flux du passif</u>

Cette étape consiste dans la préparation des flux indexés (en fonction de la politique des prestations du régime) ainsi que de la projection de l'évolution de la valeur actualisée des engagements du régime. Les opérations suivantes sont mises en place :

- L'ajustement de la fréquence initiale des flux du passif (la marge d'exploitation) à la fréquence retenue dans notre étude (trimestrielle).

- Le calcul des coefficients d'indexation des prestations à partir des taux d'inflation simulés par le GSE.

- L'indexation des flux du passif.

- Le calcul des coefficients d'actualisation des flux à partir des taux d'intérêts simulés par le GSE (en cas de flux de passif initiaux exprimés en euro réel, nous utilisons le taux réel simulé pour l'actualisation des flux futurs).

- La détermination de l'évolution des engagements actualisés à chaque instant.

<u>Etape de détermination de l'évolution du ratio de couverture</u>

Pour une allocation donnée, nous passons par les opérations suivantes pour déterminer la distribution des ratios de couverture (que ce soit le ratio de solvabilité ou le ratio de viabilité comme définies ci-dessus) :

- L'ajustement de la valeur de la réserve à chaque fin de période par rapport aux flux indexés :

Réserve fin de période = Réserve début de période capitalisée sur le marché financier + flux fin de période.

▪ La détermination de l'évolution du ratio de couverture tout au long de la période de projection et pour les différentes trajectoires simulées.

Schéma opérationnel illustrant l'approche retenue pour l'ALM

- Nous construisons initialement un très grand nombre d'allocations « tests » (ex. 100 000)

- Parallèlement nous simulons un grand nombre de scénarios de marché (ex. 5 000) pour chaque classe constituant l'actif (taux, action, immobilier)

- Nous appliquons les scénarios de marché à chacune des allocations tests, afin de simuler les trajectoires (évolutions) correspondantes de la réserve entre la date de calcul de l'allocation et l'horizon du régime.

- Nous aurons ainsi par ex 5 000 trajectoires simulées pour chacune des 100 000 allocations tests.

- Nous effectuons enfin un « tri » parmi les allocations tests, afin d'obtenir des allocations dites « admissibles » : il s'agit des allocations associées à la meilleure probabilité que le rendement soit maximal tout en respectant les contraintes définies ci-dessus. Cette probabilité est évaluée sur la base des 5 000 trajectoires simulées.

Le graphique suivant illustre les principales étapes du modèle de gestion actif-passif considéré :

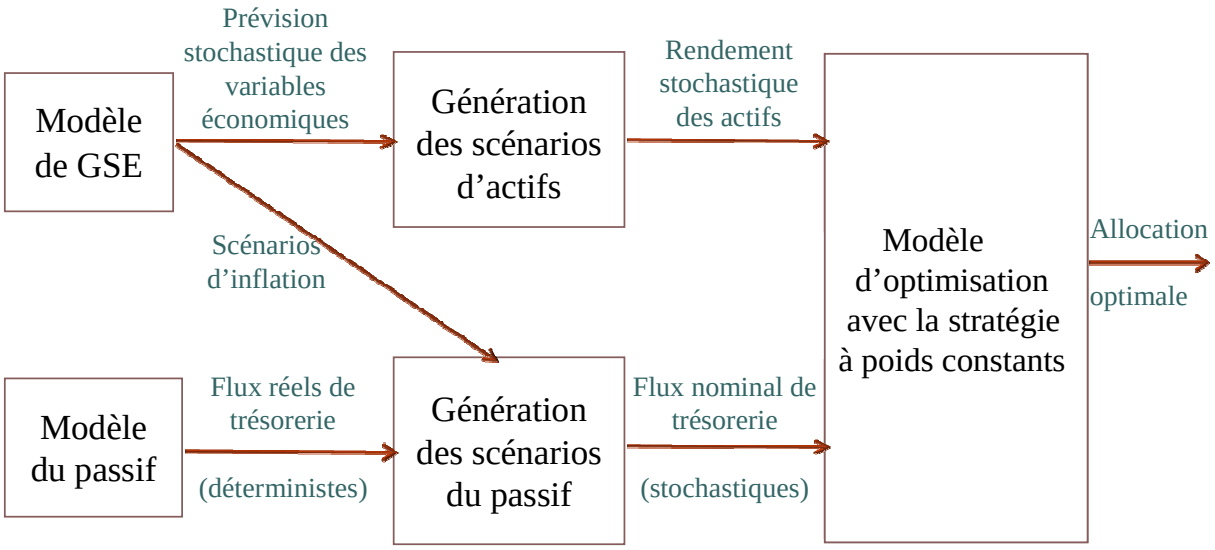


Fig. 43 : *Illustration du modèle de gestion actif-passif stochastique objet de l'application dans le cadre de la stratégie Fixed-Mix*

I-2-6 Résultats obtenus

Dans le cadre de l'application de la stratégie *Fixed-Mix*, nous optons pour la majoration des poids des différentes classes d'actifs dans l'allocation optimale de la manière suivante : pondération maximale de 40 % pour la classe des obligations à taux fixe, la classe des actions, l'immobilier et les obligations indexées et pondération maximale de 10 % pour le monétaire. L'intérêt de ce plafonnement des poids des classes d'actifs est d'ordre pratique : il permet de réduire le nombre d'allocations possibles, et donc de minorer les temps de calculs du moteur ALM.

L'objectif étant d'illustrer le caractère opérationnel du moteur ALM, la matrice d'allocation est générée en fixant un pas de maillage (de 5 %) et des bornes pour les poids de chaque actif. Les bornes ont été choisies en prenant en compte deux facteurs : le premier consiste à limiter la taille de la matrice des allocations possibles et le second consiste à couvrir un ensemble d'allocations pertinentes et vraisemblables. Les bornes retenues pour l'illustration des résultats sont donc les suivantes :

	Borne min	**Borne sup**
Monétaire	0,00 %	30,00 %
ZC 5 ans (nominal)	10,00 %	30,00 %
ZC 10 ans (nominal)	10,00 %	30,00 %
ZC 15 ans (nominal)	0,00 %	0,00 %
ZC 8 ans (réel)	10,00 %	30,00 %
ZC 15 ans (réel)	0,00 %	0,00 %
Obligation crédit	0,00 %	30,00 %
Actions	25,00 %	40,00 %
Immobilier	0,00 %	10,00 %

Tab. 29 : *Bornes retenues pour les poids de différentes classes d'actifs*

D'un autre côté, deux scénarios économiques ont été considérés : un scénario de stress et un scénario alternatif. En pratique, le but de ce type de scénario est de tester la robustesse des allocations optimales obtenues avec le scénario central (présenté dans la section III du chapitre 2 de la partie I). Ci-après les paramètres relatifs à un scénario de stress (avec forte baisse des rendements des actions : chute de 30 % sur un an) et au scénario alternatif (baisse continue de l'inflation moyenne de 100 pb). Il est rappelé que les paramètres utilisés pour les différents scénarios sont purement illustratifs.

Paramètre	Scénario central	Scénario stress (forte baisse des rendements des actions -30%)	Scénario alternatif (baisse de l'inflation moyenne de 100 bp)
Inflation à long terme			
Vitesse de retour à la moyenne	**0,01%**	**0,01%**	**0,01%**
Volatilité	**5,00%**	**5,00%**	**5,00%**
Moyenne	**2,00%**	**2,00%**	1,00%
Valeur initiale	**2,00%**	**2,00%**	**2,00%**
Inflation à court terme			
Vitesse de retour à la moyenne	**10,00%**	**10,00%**	**10,00%**
Volatilité	**1,00%**	**1,00%**	**1,00%**
Valeur initiale	**1,00%**	**1,00%**	**1,00%**
Borne inférieur	**-1,00%**	**-1,00%**	**-1,00%**
Borne supérieur	**3,00%**	**3,00%**	**3,00%**
Taux d'interêt réel long			
Vitesse de retour à la moyenne	**35,00%**	**35,00%**	**35,00%**
Volatilité	**2,50%**	**2,50%**	**2,50%**
Moyenne	**3,00%**	**3,00%**	**3,00%**
Valeur initiale	**2,50%**	**1,50%**	**2,50%**
Taux d'interêt réel court			
Vitesse de retour à la moyenne	**40,00%**	**40,00%**	**40,00%**
Volatilité	**2,50%**	**2,50%**	**2,50%**
Valeur initiale	**1,00%**	**1,00%**	**1,00%**
Borne inférieur pour les taux	**-1,00%**	**-1,00%**	**-1,00%**
Borne supérieur pour les taux	**4,00%**	**4,00%**	**4,00%**
Action			
Valeur initiale	**100,00%**	**100,00%**	**100,00%**
Moyenne regime 1 (faible volatilité)	**7,50%**	**chute 30% sur un an**	6,50%
Volatilité du régime 1 (faible volatilité)	**16,00%**	**16,00%**	**16,00%**
Moyenne régime 2 (forte volatilité)	**-10,00%**	**chute 30% sur un an**	-11,00%
Volatilité du régime 2 (forte volatilité)	**27,00%**	**27,00%**	**27,00%**
Immobilier			
Vitesse de retour à la moyenne	**15,00%**	**15,00%**	**15,00%**
Volatilité	**10,00%**	**11,00%**	**10,00%**
Moyenne	**6,00%**	**6,00%**	5,00%
Valeur initiale	**5,00%**	**5,00%**	**5,00%**
Crédit			
*Coefficient d'battement**			
Maturité 5 ans	**97,18%**	**97,18%**	**97,18%**
Maturité 4 ans et 9 mois	**97,17%**	**97,17%**	**97,17%**
Valeur initiale	-	-	-
Moyenne regime 1 (faible volatilité)	-	-	-
Volatilité du régime 1 (faible volatilité)	-	-	-
Moyenne régime 2 (forte volatilité)	-	-	-
Volatilité du régime 2 (forte volatilité)	-	-	-
Monétaire (lognormal)			
Moyenne	**3,50%**	**3,50%**	2,50%
Volatilité	**1,50%**	**1,50%**	**1,50%**

Tab.30 : *Paramètres retenus pour les différents scénarios*

Il est à préciser que le TRI et le médian de la réserve sont calculés sur un horizon égal à l'horizon d'allocation (9 ans). Par ailleurs, la VaR du ratio de viabilité et la VaR du ratio de solvabilité sont établis pour un horizon d'allocation de 9 ans (sachant par ailleurs que l'horizon de la contrainte de solvabilité est de 20 ans et que l'horizon de la contrainte de viabilité est de 30 ans). La performance et la volatilité sont des valeurs moyennes calculées sur l'ensemble de la projection (50 ans).

Au niveau des résultats sur l'allocation optimale, nous observons à première vue que le poids du monétaire est nul. Aussi, nous observons que les poids cumulé des obligations nominales et l'obligation crédit est de 40 %. Cette valeur est de 10 % pour les obligations indexées à l'inflation et de 50 % pour les actifs risqués (actions et immobilier). En outre l'allocation optimale présentée ci-dessous ne comporte pas d'obligation nominale et d'obligation indexée à l'inflation de maturité 15 ans, étant donné que ces poids sont nuls dans la matrice des allocations testées.

Le tableau ci-après présente les dix meilleures allocations parmi les 2 985 allocations testées avec le scénario économique central, la meilleure allocation est marquée en bleu :

TRI moyen (9 ans)	10,631%	10,630%	10,629%	10,630%	10,629%	10,631%	10,631%	10,632%	10,630%	10,630%
Médian de la réserve (9 ans)	10133	10137	10124	10125	10136	10126	10138	10138	10125	10133
VaR Ratio de pérennité (9 ans)	1,06	1,06	1,06	1,06	1,06	1,06	1,06	1,06	1,06	1,06
VaR Ratio de financement (9 ans)	2,786	2,788	2,792	2,792	2,787	2,792	2,788	2,792	2,792	2,786
Performance moyenne à long terme (50 ans)	5,551%	5,553%	5,545%	5,546%	5,552%	5,548%	5,553%	5,549%	5,547%	5,550%
Volatilité à long terme (50 ans)	7,396%	7,414%	7,385%	7,370%	7,382%	7,373%	7,381%	7,375%	7,369%	7,370%
Monétaire	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
ZC 5 ans (nominal)	10%	15%	30%	25%	20%	15%	10%	10%	20%	15%
ZC 10 ans (nominal)	15%	20%	10%	10%	20%	10%	20%	10%	10%	15%
ZC 15 ans (nominal)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
ZC 8 ans (réel)	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%
ZC 15 ans (réel)	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Obligation crédit	15%	5%	0%	5%	0%	15%	10%	20%	10%	10%
Actions	40%	40%	40%	40%	40%	40%	40%	40%	40%	40%
Immobilier	10%	10%	10%	10%	10%	10%	10%	10%	10%	10%

Tab.31 : *Liste des dix meilleures allocations obtenues dans l'exemple illustratif*

Nous notons que le poids du monétaire est nul dans les 10 allocations. Nous observons également sur l'ensemble des allocations une répartition équilibrée entre les actifs risqués (actions et immobilier), qui ont un poids total de 50 %, et les actifs moins risqués (obligations), qui ont également un poids total de 50 %.

En complément, un test est mené sur les 10 meilleures allocations obtenues. Le test sur ces dix allocations est réalisé pour s'assurer de leur robustesse en cas de scénarios de stress (forte baisse des actions sur un an) ou alternatif (diminution du niveau d'inflation de long terme) : il s'agit ainsi de s'assurer que ces 10 meilleures allocations vérifient les contraintes dans les deux scénarios complémentaires retenus (scénario de stress et scénario alternatif). En pratique, cela consiste à lancer la partie du code relatif aux contraintes considérées en retenant en *input* la matrice des 10 meilleures allocations.

Dans le cadre de ce test, il s'avère que les 10 allocations présentées comme les meilleures dans le scénario central vérifient bien les contraintes dans les deux scénarios complémentaires. Aussi, les allocations optimales proposées semblent résister à des scénarios

de projections des actifs différents de ceux retenus dans le scénario central. La conclusion de ce test confirme le caractère robuste des dix allocations testées.

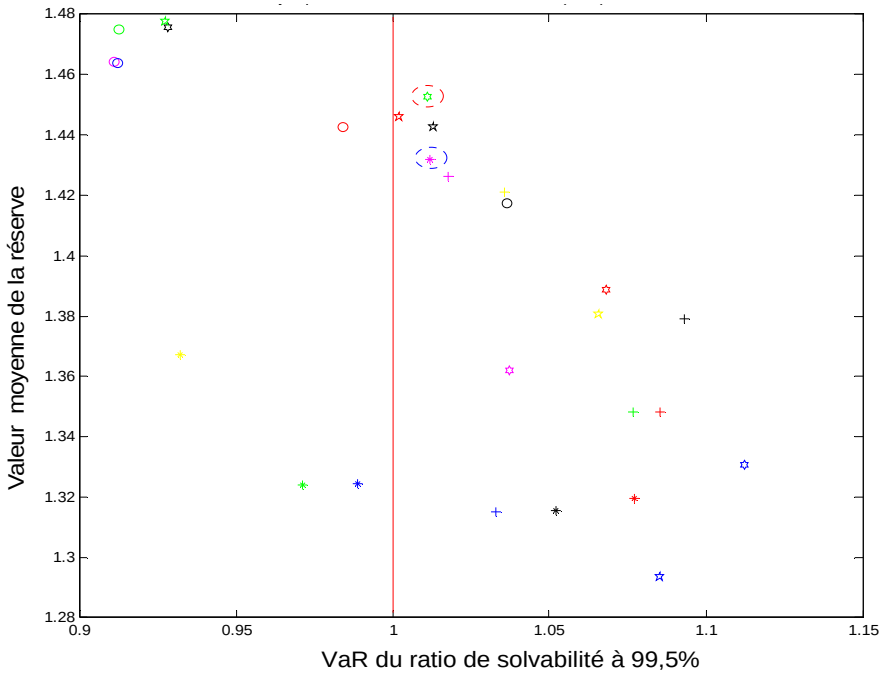


Fig. 44 : *Ratio de couverture des flux jusqu'en 2029 sur un horizon de 9 ans (2018) et un niveau de confiance de 97,5 %*

Les résultats suivants sont indiqués à titre d'illustration : Valeur à Risque (VaR) du ratio de solvabilité à 99,5 % et valeur moyenne de la réserve dans le graphique 44, ceci pour un échantillon donné d'allocation. On obtient alors une courbe d'efficience au sens de la théorie de Markowitz [1952] : un certaines allocations permettent d'avoir plus de rendement (valeur moyenne de la réserve) pour un niveau de risque plus réduit (VaR du ratio à 99,5 %). Il s'agit du cas, par exemple, des deux allocations ayant le couple de rendement/risque encerclé dans le graphique 44 et qui garantissent un niveau de VaR supérieur à 1.

Le graphique 45 illustre, quant à lui, un test effectué sur la sensibilité de l'allocation optimale par rapport aux changements des niveaux des rendements espérés des actions.

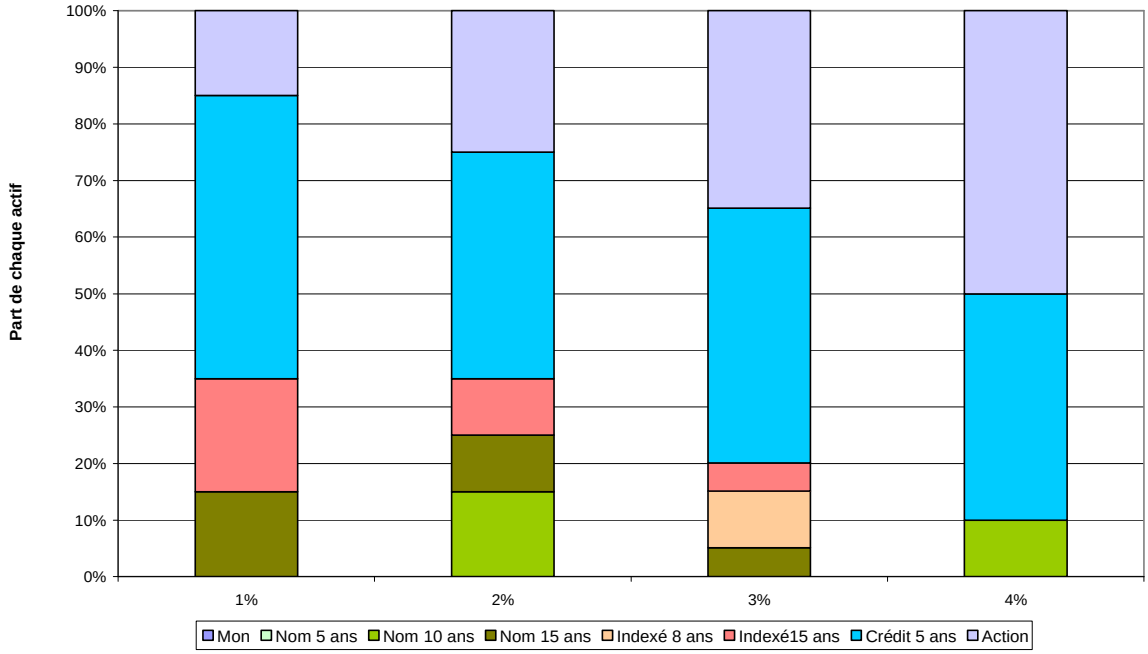


Fig. 45 : *Sensibilité de l'allocation optimale par rapport à l'espérance de rendement des actions*

Deux constats peuvent être avancés à ce titre :

- une sensibilité significative de l'allocation stratégique d'actifs par rapport aux hypothèses de départ sur les rendements des actions.
- une sensibilité croissante : plus les hypothèses de rendement sont élevées plus la variation des poids de l'allocation stratégique optimale est significative dans notre cas.

Dans ce qui suit nous nous intéressons à l'impact du recours à deux matrices de corrélation différentes (cf. section I, chapitre 2 de la partie I) entre les actions et les autres classes d'actifs : une première matrice reflétant la structure dépendance dans un contexte « normal » et une deuxième matrice reflétant une situation de « crise » (appelée matrice de « corrélation à risque »).

Pour cela, un premier tableau (cf. tableau 32) est présenté indiquant l'allocation optimale obtenue dans le cas d'un seul régime (1 matrice de corrélation) et dans le cas de deux régimes (2 matrices de corrélation), toutes choses étant égales par ailleurs. Nous constatons que, dans ce dernier cas, nous sommes moins exposés aux classes d'actifs risqués (exposition totale aux actions et à l'immobilier à la hauteur de 65 % contre 70 % dans le cas d'un seul régime. Cela constitue une réponse à une aversion au risque supplémentaire prise en compte via la matrice de « corrélation à risque »).

	Alloc optim 1 matrice de corr	Alloc optim 2 matrice de corr
Nom 5 ans	0%	0%
Nom 10 ans	10%	15%
Nom 15 ans	5%	10%
Indexées 8 ans	0%	0%
Indexées 15 ans	15%	10%
Crédit 5 ans	40%	40%
Actions	30%	25%

Tab. 32 : *Allocations optimales en fonction du nombre de régimes de corrélations*

Le même travail a été effectué au niveau des VaR (Valeur à risque) correspondant au montant de la réserve totale observée à une date donnée (horizon de l'allocation). Ainsi, pour deux niveaux de confiance possibles (97,5 % et 99,5 %), nous avons calculé ces valeurs dans le cas où nous avons un seul régime de corrélation et dans le cas où nous avons deux régimes de corrélation (cf. tableau 33). Les résultats confirment une moindre exposition dans le cas de deux régimes et montrent une sensibilité significative du niveau de la VaR (en montant) par rapport à l'approche retenue (dans le cas du seuil de 97,5 %, nous remarquons bien un doublement du montant extrême en passant de l'approche à un seul régime à l'approche à deux régimes de corrélation).

	VaR 97,5 %	VaR 99,5 %
Alloc optim 1 matrice de corr	203	-3740
Alloc optim 2 matrice de corr	435	-3255

Tab.33 : *Valeur à Risque de la valeur de la réserve en fonction du nombre de régimes de corrélations*

L'histogramme ci-dessous confirme également ce constat via des queues extrêmes relativement plus épaisses (et plus étendues vers la gauche) dans le cas d'un seul régime par rapport à ceux observés dans le cas de deux régimes.

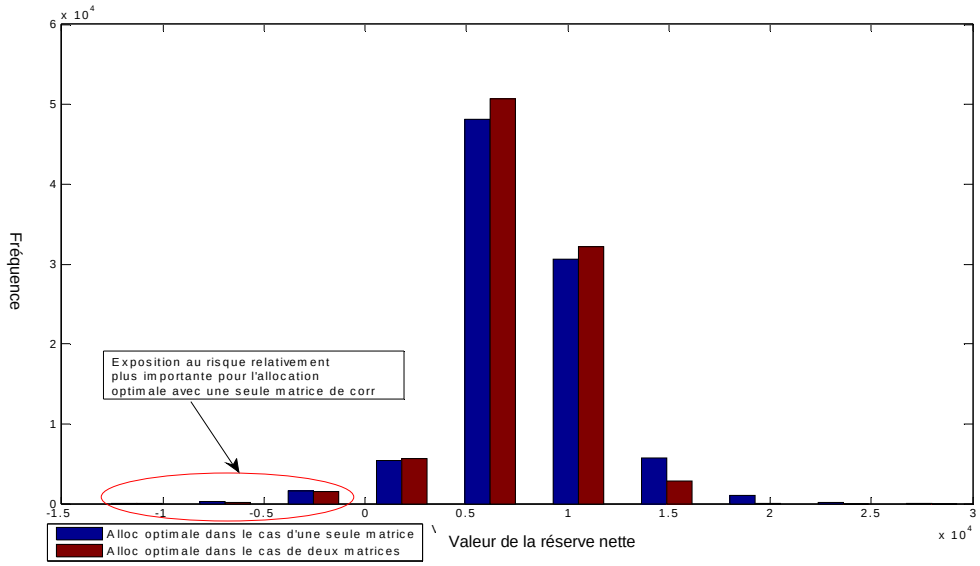


Fig. 46 : *Distribution de la valeur de la réserve totale (nette des flux de trésorerie) dans le cas de projection de scénarios avec deux matrices de corrélation*

II- Proposition de méthodes numériques

La résolution du problème posé peut être présentée en trois étapes. La première consiste à générer des trajectoires pour chaque classe d'actifs. La deuxième consiste à projeter chaque allocation en fonction de la règle de gestion du dispositif et à en retenir une, qui conduit à une trajectoire de rendement de l'actif synthétique. La troisième étape vise à déterminer la valeur de la fonction objective (le passif étant déterministe). Les allocations susceptibles d'être retenues seront alors celles qui maximisent la fonction objective et respectent les contraintes du dispositif.

Ce paragraphe présente les méthodes de mise en œuvre opérationnelle de l'allocation stratégique.

Le nombre d'allocations possibles avec les neufs classes (notamment lorsque le pas de l'allocation est égal à 1 %) risqueraient d'engendrer des temps de calcul importants, ce qui limiteraient les performances opérationnelles du modèle. Dans ce contexte, la réduction du nombre de classes ou l'augmentation du pas dans le choix des allocations possibles amélioreraient les performances opérationnelles du modèle. Néanmoins, la réduction du nombre de classes d'actifs ne peut aller au-delà d'un minimum de 2 classes d'actifs sur les taux nominaux (hors monétaire) et les taux réels, à savoir une maturité moyen terme et une maturité long terme. Cela se justifie par plusieurs raisons :

- pouvoir tester dans l'allocation, au moins l'échéance de réinvestissement et de « refixation » du rendement en regard de la couverture à l'inflation ;

- les engagements des régimes pouvant être très longs, il convient de tenir compte un minimum du positionnement sur la courbe des taux.

Quelle que soit la méthode retenue, le nombre d'allocations possibles nécessite de recourir à des algorithmes d'optimisation efficaces, l'approche consistant à tester de manière exhaustive l'ensemble des allocations possibles n'étant pas envisageable.

Cette partie traite donc un problème d'optimisation dans le cas où la fonction objectif est estimée à l'aide de simulations. Il présente un algorithme d'optimisation globale d'une fonction non convexe et bruitée. L'algorithme est construit après une étude de critères de compromis entre, d'une part, l'exploration de la fonction objectif en de nouveaux points et d'autre part l'amélioration de la connaissance de celle-ci, par l'augmentation du nombre de tirages en des points déjà explorés. Une application numérique illustre la conformité du comportement de cet algorithme à celui prévu théoriquement. Les performances de l'algorithme sont analysées au moyen de différents critères de convergence. Des zones de confiance sont également proposées. Enfin, une comparaison avec un algorithme classique d'optimisation stochastique est menée.

II-1 Introduction

II-1-1 Le problème d'optimisation

Nous allons chercher à déterminer l'optimum global d'une fonction réelle définie sur un ensemble Θ soumise à un bruit, ainsi que les paramètres de Θ conduisant à cet optimum. Ainsi, dans le contexte de l'allocation stratégique optimale d'actifs, une allocation

particulière, $\theta \in \Re^d$, $d \in \mathrm{N}^*$ peut nécessiter de simuler de nombreuses trajectoires de plusieurs actifs, sur une succession de périodes, afin de déterminer un unique indicateur de risque ou de gain, bruité, $F(\theta)$: l'indicateur en question peut être un ratio de financement, un indicateur de gain pénalisé en fonction du risque, ou un indicateur synthétique de compromis risque/gain. Le calcul de chaque $F(\theta)$ est coûteux en terme de temps de calcul, et les allocations conduisant à un optimum de $f(\theta) = E[F(\theta)]$ sont recherchées pour l'allocation optimale.

Nous considérons ici une fonction objectif réelle $f(\theta)$ d'un paramètre $\theta \in \Theta$, $\Theta \subset \mathrm{R}^d$, $d \in \mathrm{N}^*$:

$$f : \Theta \to \mathrm{R}, \quad \Theta \subset \mathrm{R}^d$$

Nous supposerons que f est continue et bornée, non nécessairement dérivable, et que Θ est une union finie de *d*-simplexes, typiquement un ensemble de pourcentages d'allocation possibles, inclus dans $[0,1]^d$. Enfin, f n'est pas nécessairement une fonction convexe de θ, de sorte que la fonction peut posséder plusieurs optima locaux.

En outre, nous supposerons que f n'est pas directement connue, mais qu'elle est estimée à l'aide de simulations. En tout point $\theta \in \Re^d$, du fait des erreurs d'estimation, l'observateur ne peut accéder qu'à des réalisations d'une variable aléatoire $F(\theta) = f(\theta) + \varepsilon(\theta)$, où $\varepsilon(\theta)$ représente un bruit d'espérance nulle, dont nous postulerons l'existence d'une variance finie, non nécessairement homogène en θ. Nous postulons que les $\{\varepsilon(\theta)\}_{\theta \in \Theta}$ sont mutuellement indépendants. La fonction bruitée F est donc telle que :

$$\forall \theta \in \Theta : \begin{cases} E[F(\theta)] = f(\theta) \\ V[F(\theta)] < \infty \end{cases}$$

Nous nous placerons dans le cadre d'une minimisation. Sous ces hypothèses, chercher à optimiser la fonction bornée $f(\theta) = E[F(\theta)]$ revient alors à rechercher, à partir de réalisations ponctuelles indépendantes de la fonction bruitée F :

1. L'unique valeur minimale de la fonction objectif f,

$$m^* = \inf_{\theta \in \Theta} E[F(\theta)]$$

2. L'ensemble des paramètres conduisant à une valeur proche de m^* :

$$S_x = \{\theta \in \Theta, E[F(\theta)] \leq x\}$$

pour tout réel x donné dans un voisinage de m^*

Le premier point est intéressant, mais s'agissant d'allocation d'actifs, le résultat le plus utile est naturellement le second. En un mot, nous cherchons tous les paramètres θ conduisant à un $f(\theta)$ proche de l'optimum.

En présence d'une incertitude sur l'estimateur de f, la recherche d'un paramètre θ^* conduisant à un optimum global supposé de f nécessite l'exploration de tous les paramètres susceptibles de conduire à un optimum global inférieur. Il sera ainsi nécessaire de connaître l'ensemble S_x des paramètres conduisant à une valeur estimée de f proche de m^*.

En outre, en présence de plusieurs paramètres solutions, seul l'utilisateur de l'algorithme peut décider lequel privilégier. C'est la raison pour laquelle nous rechercherons l'ensemble des paramètres solution, et non un seul point de cet ensemble.

Lorsque la fonction F est déterministe, dans le cas où $f(\theta) = F(\theta)$, différentes méthodes d'optimisation peuvent être proposées, comme les méthodes de descente de gradient, de Newton-Raphson, de Hooke et Jeeves, la méthode de Nelder, Mead [1965], ou des méthodes spécifiquement adaptées à certaines formes de f, comme lorsque f est convexe. Ces méthodes ne garantissent toutefois pas que l'optimum obtenu est un optimum global. S'agissant d'optimisation globale déterministe, les premières recherches datent d'environ trente ans et sont attribuées à Hansen [1979]. L'optimisation Lipschitzienne, l'algorithme de Schubert (cf. Schubert [1972]) ou l'algorithme DIRECT (cf. Jones et al. [1993]) sont des méthodes largement utilisées dans un cadre déterministe. D'une façon plus générale, ces algorithmes peuvent s'intégrer dans le cadre d'algorithmes de type *Branch and Bound* (cf. Lawler et al. [1966]), où la zone à explorer (ici Θ) est partitionnée en plusieurs zones (*branching*), dont certaines sont exclues de l'analyse selon certains critères (*bounding*).

Les critères d'exclusions de ce type d'algorithme reposent notamment sur des propriétés de la fonction objectif et sur une arithmétique d'intervalle (cf. Wolfe [1996]), nous parlons alors d'optimisation globale déterministe, permettant de garantir avec certitude l'absence d'optimum sur les zones exclues. En considérant que f est le résultat d'une expérience, généralement coûteuse en terme de temps de calcul, les techniques d'optimisation entrent dans le cadre de *Computer Experiments* et la représentation de la fonction en dehors des points d'observation peut s'appuyer sur différentes techniques de régression, de Krigeage, ou de champs gaussien (cf. Krige [1951], Jones et al. [1998], Santner et al. [2003]). Une revue des différentes techniques utilisées dans le champ de l'optimisation globale pourra être trouvée dans Horst, Parlados [1995], ainsi que dans plusieurs thèses récentes (cf. Emmerich [2005], Ginsbourger [2009] et Villemonteix [2009]).

Lorsque la fonction F est aléatoire, lors de la recherche d'un unique optimum non nécessairement global, une très large littérature existe sur les algorithmes stochastiques pour la recherche d'optimum. Lorsque $d = 1$ des algorithmes tels que celui de Kiefer-Wolfowitz (voir Kiefer, Wolfowitz [1952] peuvent être utilisés. Dans le cas $d > 1$, une extension due à Blum peut être exploitée (voir Blum [1954]). Strugarek [2006] présente un ensemble d'algorithmes plus récents et plus détaillés portant sur l'optimisation stochastique. S'agissant de la recherche de racines, que nous pourrons également traiter avec l'algorithme ici présenté, des techniques classiques sont celles dérivées de l'algorithme de Robbins-Monro (Robbins, Monro [1951]). De même, l'algorithme de Cohen-Culioli permet de résoudre des problèmes très proches avec une contrainte de contrôle de la probabilité de ruine (voir Cohen et al. [1994]). Nous ne détaillerons pas ici l'ensemble des techniques utilisées dans les champs de l'optimisation stochastique (recherche d'optimum), ou de l'approximation stochastique (recherche de racines).

Enfin, lorsque la fonction F est aléatoire, et lorsque nous cherchons un minimum global d'une fonction non nécessairement convexe, le bruit complique encore le problème difficile de l'optimisation globale (voir Bulger et al. [2005] pour une discussion sur ce sujet). Certaines méthodes stochastiques ont été adaptées à un environnement bruité : nous pouvons citer les méthodes génétiques (voir Alliot [1996], Mathias et al. [1996] en présence de bruit), ainsi que les méthodes de recuit simulé (voir Aarts et al. [1985], Brankeet al. [2008] en présence de bruit). Encore peu connues dans le domaine de l'actuariat, quelques extensions

d'algorithmes de type *Branch and Bound* sont également proposées dans un cadre stochastique (cf. par exemple Norkin et al. [1996]).

Le risque de se tromper d'optimum Lorsque F est aléatoire, nous pouvons envisager l'usage d'algorithmes stochastiques d'optimisation locale. Toutefois, tout comme les méthodes déterministes de descente de gradient, ces algorithmes requièrent de choisir un point de départ suffisamment proche d'un optimum global, et d'éviter les plus évidents optima locaux. Sauf pour quelques fonctions f particulières, il s'agit alors d'appréhender dans un premier temps la forme de la fonction f. Une exploration globale de la fonction s'avère donc souvent nécessaire, le risque étant moins de mal estimer la valeur d'un optimum global, que de se tromper d'optimum en choisissant indûment un mauvais optimum local.

Une seconde piste serait d'établir un grand nombre de tirages de F, afin de se ramener, à une marge d'erreur près, au cas déterministe. Cela autoriserait l'usage d'algorithmes d'optimisation locale déterministe, généralement rapides. Toutefois, la recherche d'un optimum local se ferait au prix d'une exploration préalable de la fonction. De surcroît, l'algorithme utilisé serait lourdement pénalisé par le grand nombre de tirages requis pour F.

De la même façon, cette pénalité frapperait les algorithmes d'optimisation globale déterministe. Enfin, les méthodes génétiques ou le recuit simulé ne visent pas directement à proposer des zones de confiance pour les optimiseurs de la fonction objectif. Il nous a semblé délicat de quantifier, avec ces méthodes, le risque de proposer un optimum local et d'ignorer un optimum global meilleur.

Explorer ou connaître A titre illustratif, plaçons nous un instant dans un cadre simplifié, en présence d'un unique minimiseur $\theta^* = arg\,min_{\theta\in\Theta} f(\theta)$. Supposons que l'observateur puisse réaliser n tirages de F en chaque point d'un ensemble $\Theta_m = \{\theta_1,...,\theta_m\}$. En supposant donnée la valeur n, qui détermine la précision de la connaissance de f, et Θ_m qui détermine l'ampleur de l'exploration de f, un estimateur envisageable $\hat{\theta}^*$ de l'unique minimiseur θ^* est le suivant :

$$\hat{f}_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} F_i(\theta) \qquad \theta \in \Theta_m$$

$$\hat{\theta}^*_{n,m} = arg \min_{\theta\in\Theta_m} \hat{f}_n(\theta)$$

La question se pose ici de l'arbitrage entre le choix d'un n élevé ou d'un cardinal de Θ_m élevé. Les algorithmes usuels d'optimisation supposent généralement que l'ensemble des points $F(\theta)$ sont connus, et ignorent, à notre connaissance, la question pratique de l'optimisation du couple de paramètre (n, m). En effet, la détermination d'une réalisation de la variable aléatoire F peut prendre un temps de calcul important. En conséquence, n réalisations de F en chaque point de Θ_m conduisent à un total de $(n.m)$ tirages de F. Or, les tirages étant coûteux en termes de temps de calcul, le nombre de tirages est nécessairement limité. Pour un budget de tirages fixé, le choix d'un nombre de simulations n élevé réduit la variance de $\hat{f}_n$, et donc améliore ponctuellement la connaissance de f, mais limite également l'exploration de la fonction en d'autres points.

Cette illustration simplifiée introduit la problématique de l'arbitrage entre l'exploration de la fonction en différents points et la connaissance ponctuelle de celle-ci. Trouver les minimiseurs de f va requérir d'une part d'explorer la fonction en différents points (nous parlerons d'exploration), et d'autre part d'opérer également plusieurs tirages de F en chaque point exploré pour obtenir un estimateur non biaisé et aussi peu dispersé que possible de l'espérance de la fonction (nous parlerons de connaissance ponctuelle). Nous proposerons dans cette partie un algorithme qui vise à répondre à cet arbitrage.

Objectif poursuivi L'objectif que nous poursuivrons ici est l'exploration des zones susceptibles de contenir un minimum global de f : nous chercherons donc à explorer sélectivement la fonction f, de façon à avoir une vision globale de celle-ci, tout en privilégiant certaines zones d'intérêts, comme les minima globaux, dans un souci d'économie du nombre de tirages de la fonction F. L'algorithme vise au final à estimer m^* ainsi que S_x, pour x donné dans un voisinage de m^*.

II-1-2 Approche retenue

L'approche proposée ici s'apparente à une approche de type *Branch and Bound*, mais aucune zone n'est jamais définitivement exclue de l'analyse, l'idée étant ici d'ordonner les zones en fonction de la probabilité qu'une zone contienne un optimum, selon un modèle que nous détaillerons.

Nous nous placerons dans le cadre de la minimisation d'une fonction, nous chercherons ici à quantifier la probabilité qu'une zone contienne un minimum plus petit qu'un minimum observé, pour au final construire l'ensemble des zones susceptibles de contenir un minimum global.

Grille à pas fixe Une solution simple et très commune est l'utilisation d'une grille de simulation, à pas fixe. Selon cette solution, la fonction f est explorée sur un ensemble Θ_m. Θ_m est une grille de pas δ, toutes les composantes de $\theta \in \Theta_m$ parcourent l'ensemble des valeurs multiples de δ dans Θ : $\Theta_m = \left\{ \theta \in \Theta, \frac{1}{\delta}\theta \in \mathrm{N}^d \right\}$. Le nombre n de tirages de $F(\theta)$ est identique pour chaque $\theta \in \Theta_m$. Cette solution est néanmoins très onéreuse en terme de temps de calcul, dans la mesure où n simulations seront conduites sur chacun des points, y compris ceux très éloignés d'un optimum, pour lesquels la meilleure connaissance de f n'apporte quasiment rien.

Grille à pas variable Nous chercherons donc à mettre en place une grille à pas variable, où les retirages de F se feront principalement dans les zones susceptibles d'accueillir le minimum. Il va s'agir d'une part d'estimer f en de nouveaux points θ (sommets), et d'autre part de répartir de nouvelles simulations entre les anciens sommets et les nouveaux, en fonction de l'intérêt que peut avoir, sur la connaissance du minimum global, l'ajout de simulations en chacun de ces points. Les positions des sommets envisagés ne seront plus régulièrement réparties, et le nombre de simulations conduites en chaque sommet va différer selon les sommets. Une première difficulté est le choix d'une forme convenable pour les zones de recherche, pour les cellules de la grille. Une autre difficulté est que nous avons besoin de deviner où conduire les futures simulations. Pour cela, il faut avoir une idée de comment va

évoluer la fonction entre les points explorés : il faut en un sens deviner quel pourra être l'impact de futures simulations avant même de les réaliser.

Algorithmes présentés et structure du document Nous allons aborder dans cette étude deux algorithmes réalisant une grille à pas variable. Dans la sous-section II-2, nous envisagerons le cas où le nombre de tirages en chaque point est fixe, l'algorithme cherchant alors uniquement à déterminer les prochains points où évaluer la fonction F . Dans la sous-section II-3, nous introduirons la possibilité d'opérer des retirages en des points déjà explorés. Enfin, dans une dernière sous-section II-4, nous présenterons des illustrations simples et visuelles du comportement de l'algorithme proposé sur une fonction bimodale élémentaire. Une comparaison avec un algorithme classique d'optimisation stochastique sera également menée.

II-2 Une grille à pas variable par subdivision systématique

II-2-1 Forme des zones de recherche

Zone initiale de recherche En pratique, par exemple lors de l'optimisation de choix d'investissements de d +1 actifs, les proportions investies dans chacun des actifs se somment à un. Il suffit alors de rechercher d pourcentages d'allocation, l'allocation numéro d +1 se déduisant des d allocations précédentes. Il s'agit donc d'optimiser une fonction de R^d dans R .

Nous nommerons $Z_0 \subset R^d$ la zone initiale de recherche, simplexe standard orthogonal constitué des sommets (1, 0, …, 0), (0, 1, 0,…, 0), ... , (0,…, 0, 1), ainsi que du sommet (0,…, 0).

Cela correspondra bien à la situation où la seule contrainte est d'avoir une somme des composantes inférieure à un, par exemple d pourcentages d'allocation d'actifs dont la somme est inférieure à 100 %, la différence avec 1 formant l'allocation de l'actif numéro d +1 :

$$Z_0 = \{(x_1,...,x_d) \in R^d, x_1 + ... + x_d \leq 1, x_1 \geq 0, ...x_d \geq 0\}$$

Dans les situations où l'optimum est à rechercher sur une zone plus complexe soumise à de nombreuses contraintes, l'algorithme proposé sera applicable si cette zone peut être représentée dès l'initialisation par une union finie de simplexes. Dans tous les cas, nous supposerons que l'ensemble des sommets de la ou des zones initiales ont été explorés, par la réalisation de tirages de $F(\theta)$ en chacun des sommets de l'enveloppe de ces zones.

Mécanisme de scission L'idée d'une grille à pas variable est de séparer la zone de recherche de l'optimum en plusieurs zones de différentes tailles. A chaque étape, la scission d'une zone peut se faire en opérant des tirages de F en un ou plusieurs points non encore explorés de Θ .

Certains algorithmes d'optimisation globale fonctionnent, dans un contexte particulier déterministe, par subdivision de la zone à explorer. Nous pouvons notamment citer l'algorithme DIRECT (pour *DIviding RECtangles*, cf. Jones et al. [1993]), qui subdivise un pavé de R^d en plusieurs sous-pavés. Si la subdivision d'une zone nécessite l'exploration de n^+ nouveaux points, il paraît préférable de choisir $n^+ = 1$. Dans ce seul cas, le choix d'un nouveau point de subdivision se fait alors en connaissance de tous les précédents tirages de la fonction. Nous avons ici choisi un mode de division qui d'une part nous semblait plus adapté

aux problèmes définis sur une union de simplexes, et d'autre part ne nécessitait l'exploration que d'un unique point à chaque subdivision.

Nous supposerons que les zones de recherche sont des ensembles convexes, de façon à faciliter d'éventuelles interpolations au coeur de chaque zone. Imaginons un pavé de $\Re^d$ dont on connaît les sommets, et à l'intérieur duquel nous ajouterions un point θ_c. Comment découper rapidement ce pavé en une partition d'ensembles convexes dont l'enveloppe convexe contiendrait θ_c ? La réponse n'étant pas si évidente, nous opterons pour le choix suivant :

- Les zones considérées seront délimitées par d +1 sommets.
- Chaque nouveau point sera ajouté sur un segment de l'enveloppe convexe de la zone.
- A chaque étape, chaque zone sera éventuellement scindée en deux ensembles convexes.

Tout point à l'intérieur d'une zone appartiendra à une zone convexe délimitée par d +1 points, et le choix d'un nombre de sommets égal à d +1 facilitera par la suite la séparation d'une zone en plusieurs zones. D'autre part, le choix d'un nombre fixe de sommets délimitant chaque zone sera de nature à faciliter l'implémentation de l'algorithme.

Nous appellerons zone un d-simplexe, c'est-à-dire un ensemble convexe inclus dans Θ, avec $\Theta \subset \mathrm{R}^{\mathrm{d}}$, dont l'enveloppe convexe est déterminée par d +1 sommets distincts (les sommets formant un repère affine de $\Re^d$). Par la suite, nous noterons $S(Z)$ l'ensemble des d +1 sommets délimitant une zone convexe Z. La zone initiale de recherche de l'optimum est notée Z_0. A l'issue de l'étape numéro k, la zone de recherche est subdivisée en un ensemble de zones recouvrant Z_0, dont l'intersection est de mesure nulle. Cet ensemble de zones est noté Z_k.

Considérons un ensemble de sommets E délimitant une zone Z et deux points distincts θ_1 et θ_2 de cet ensemble. Supposons que θ_c soit le barycentre (par exemple équipondéré) entre ces deux points. Lorsque la décision est prise de scinder cette zone en deux zones Z_1 et Z_2, les ensembles de sommets délimitant ces deux zones seront :

$$E_1 = (E \mid \theta_1) \cup \{\theta_c\}$$
$$E_2 = (E \mid \theta_2) \cup \{\theta_c\}$$

Une preuve de ce lemme est donnée dans Rullière et al. [2010]. Par souci de simplicité, nous nommerons par la suite sommets d'une zone Z l'ensemble des d + 1 sommets délimitant l'enveloppe convexe de la zone Z.

Le graphique 47 illustre en dimension 2 la façon dont une zone peut être progressivement subdivisée. Nous pouvons remarquer que le choix d'une décomposition de l'ensemble Θ en un ensemble de zones contenant des points de tirage $\{\theta_i\}_{i=1,2,\ldots,n}$ est un problème de triangulation, classique en géométrie algorithmique. Une technique de triangulation très connue est la triangulation de Delaunay (cf. De Berg et al. [2008]). Le problème étant ici de choisir un point de tirage au sein d'une zone et non pas de construire des zones pour séparer des points de tirage existants, les zones seront scindées par la simple bissection présentée.

Un des avantages de la bissection présentée est de permettre l'exploration des frontières de la zone initiale. D'autres choix de subdivision de zones sont naturellement possibles, par exemple autour de l'isobarycentre d'une zone. Comme nous pouvons l'observer sur le graphique 47 ainsi que sur les différentes illustrations de la sous-section II-4, l'union des enveloppes convexes des zones scindées forme un ensemble beaucoup plus vaste que l'enveloppe convexe du simplexe initial, de sorte que les points de tirage ne sont pas condamnés à rester dans l'enveloppe de la zone initiale.

Dans le cas de barycentres équipondérés, nous pouvons d'ailleurs montrer par récurrence qu'il est possible d'atteindre en un nombre fini d'étapes tout point de Θ de coordonnées :

$\left(\frac{i_1}{2^{k_1}},..., \frac{i_d}{2^{k_d}} \right)$ avec pour tout $j \in \{1,...,d\}, k_j \in \mathrm{N}, i_j \in \mathrm{N}, i_j \leq 2^{k_j}$

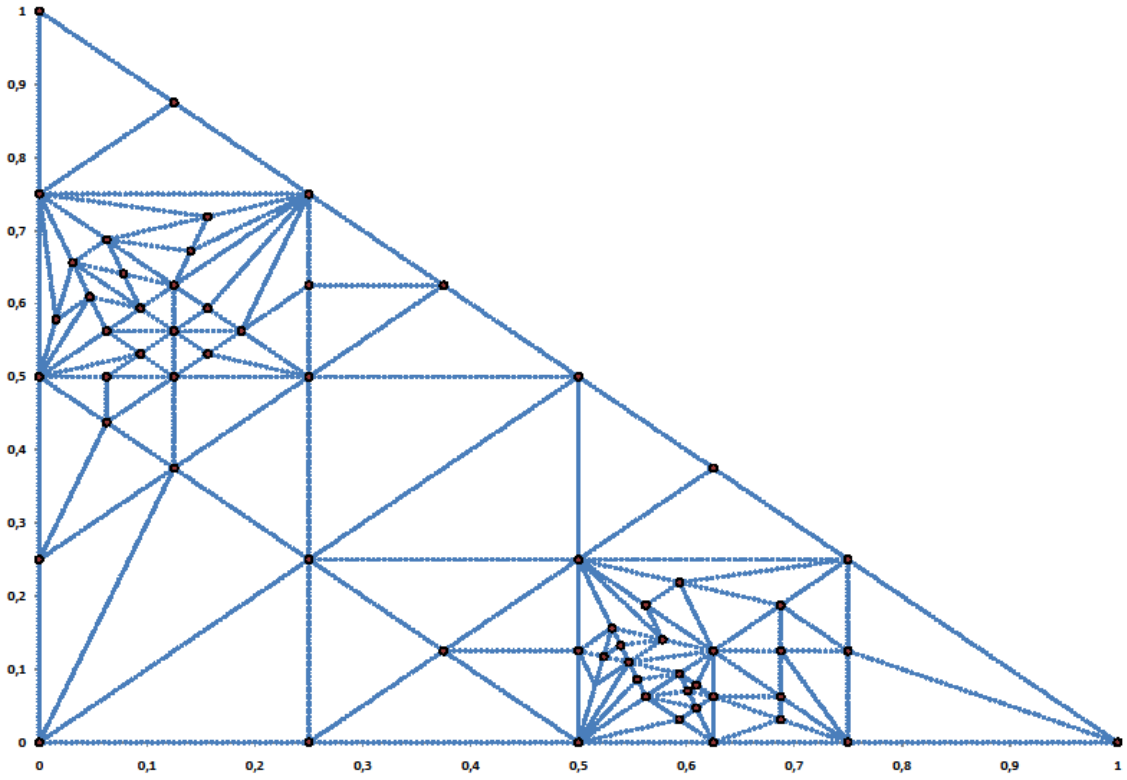


Fig. 47 : *Un exemple de partition d'une zone dans* $\Re^2$ *en 60 zones. (*d *= 2)*

II-2-2 Choix de la zone à explorer ou segmenter

Idée du potentiel d'une zone L'étude de la possibilité pour f d'atteindre un minimum global sur une zone non explorée nécessite de fixer des hypothèses : si f est supposée extrêmement erratique, f pourra franchir un seuil inférieur sur à peu près n'importe quelle zone, et les tirages opérés de F n'apporteront que très peu d'information. Une solution classiquement retenue, dans le domaine de l'optimisation globale déterministe, est le choix d'une forme Lipschitzienne pour f (cf Jones et al. [1993]) : cela revient à dire, dans un cadre déterministe,

que f pourra atteindre un minimum global sur une zone si la pente nécessaire pour cette atteinte est inférieure à un certain seuil, ou que f ne pourra pas atteindre le minimum global.

Cette logique binaire permet d'indiquer de façon certaine si f appartient ou non à un intervalle donné, et correspond à une logique d'arithmétique d'intervalle (cf. Wolfe [1996]). En l'absence d'information très précise sur f, nous avons préféré une logique probabiliste, malgré une part de subjectivité qu'elle peut engendrer : plutôt qu'une indicatrice de franchissement possible à valeur dans $\{0,1\}$, nous allons rechercher une mesure sur $[0,1]$ quantifiant la probabilité, selon la représentation de f par l'observateur, que f franchisse le seuil sur une zone. Si cette solution introduit une nécessaire subjectivité sur le calcul de cette probabilité, elle offre néanmoins l'avantage de maintenir une hiérarchie entre différentes zones, ce qui est notamment utile lorsque beaucoup de zones sont susceptibles de contenir l'optimum, et permet de ne jamais exclure *a priori* de zone.

Comme le temps de calcul de F est supposé beaucoup plus important que le temps d'exécution de l'algorithme de choix des zones à explorer, nous choisirons à chaque étape de l'algorithme une unique zone à scinder en deux ou à explorer davantage. L'objectif étant de limiter les risques qu'une zone non explorée contienne un optimum, nous piocherons à chaque étape une zone à explorer, avec une probabilité proportionnelle à un coefficient β spécifique à la zone.

Pour chaque zone Z, le coefficient $\beta(Z)$ déterminera s'il est plausible (dans un sens que nous préciserons) que la zone contienne un minimum plus petit que le minimum estimé m^*. Afin d'approcher ce coefficient sur toute la zone, nous aurons besoin de le déterminer en un unique point. Selon un modèle probabiliste que nous préciserons, nous nommerons "potentiel" d'un point une mesure d'autant plus grande que la fonction objectif est susceptible, en ce point, d'être plus petite que le minimum global observé. Une représentation probabiliste du potentiel d'un point passe par la modélisation de la connaissance incertaine de la fonction f au vu des tirages opérés en différents points, par exemple au moyen de champs aléatoires conditionnels.

Nous présentons ici une approche plus simple partant d'une modélisation de l'incertitude en dimension 1, puis agrégeant différentes incertitudes compte tenu des sommets explorés.

<u>Potentiel d'un point en direction d'un sommet</u> Considérons un sommet θ' déjà exploré d'une zone Z, et un point θ de cette zone, distinct de θ'. Par souci de simplicité, nous étudierons ici l'évolution de la fonction f entre ces deux points θ et θ' en connaissance de la seule exploration en θ'.

Même en l'absence d'erreur d'échantillonnage au sommet θ', l'évolution de la fonction f est naturellement inconnue sur le segment $]\theta' ; \theta]$: nous parlerons d'*erreur de grille*. Nous supposerons que, à partir de la seule connaissance de f au point θ', l'observateur représente $f(\theta)$ par une variable aléatoire $\tilde{f}_{\theta'}(\theta)$, choisie de loi normale, d'espérance $f(\theta')$ et dont l'écart-type est une fonction croissante de la distance $d(\theta,\theta')$. Nous choisirons :

$$\tilde{f}_{\theta'}(\theta) - f(\theta') \approx N\left(0,\sigma_g^2\right)$$

$$\sigma_g(\theta,\theta') = \sigma_K d(\theta,\theta')^{\alpha} \qquad \sigma_K > 0,\ \alpha > 0$$

Cela revient à décrire l'incertitude maximale pesant sur $\tilde{f}$, lorsque σ_g est soumis à une condition de type Hölder (ou de type Lipschitz lorsque $\alpha = 1$).

A titre d'illustration, pour $\theta_0 \in [\theta' ; \theta]$, la représentation $\tilde{f}_{\theta'}(\theta_0) - f(\theta')$ par un mouvement brownien correspond à $\alpha = 1/2$. La représentation de la pente $\left(\tilde{f}_{\theta'}(\theta_0) - f(\theta')\right)/d(\theta_0,\theta')$ par un mouvement brownien correspond à $\alpha = 3/2$. Nous évoquerons plus en détail le choix de α et σ_K au paragraphe II-2-3, ainsi que dans la sous-section d'application numérique II-4.

En présence d'erreur d'échantillonnage, nous disposons d'un estimateur $\hat{f}(\theta')$ de $f(\theta')$. Si $\hat{f}(\theta') - f(\theta')$ est elle-même une variable aléatoire de loi normale centrée et de variance $\sigma_e^2(\theta')$, indépendante de l'erreur de grille $\tilde{f}_{\theta'}(\theta) - f(\theta')$, alors :

$$\tilde{f}_{\theta'}(\theta) - \hat{f}(\theta') \approx N\left(0, \sigma_e^2(\theta') + \sigma_g^2(\theta,\theta')\right)$$

En connaissance d'une réalisation de $\hat{f}(\theta')$, nous noterons :

$$\frac{1}{2} L(\theta,\theta') = P\left[\tilde{f}_{\theta'}(\theta) \le m^*\right]$$

où m^* représente la valeur du minimum global de $f(\theta)$.

La quantité $L(\theta,\theta')$ sera nommée potentiel directionnel du point θ en direction de θ'. Le facteur de normalisation 1/2 n'a aucune incidence sur les comparaisons des potentiels entre eux (comme tout facteur strictement positif), mais permettra par la suite à la quantité *L* de pouvoir atteindre toutes les valeurs de [0, 1], et non pas seulement [0, 1/2]. Le potentiel directionnel s'interprète alors comme la probabilité que $\tilde{f}_{\theta'}$ franchisse un minimum global connu sachant que $\tilde{f}_{\theta'}$ décroît depuis le point d'accroche θ'.

Il est possible de tenir compte de l'erreur d'estimation du minimum m^* : en supposant que l'estimateur $\hat{m}^*$ est une variable aléatoire indépendante de $\tilde{f}_{\theta'}(\theta) - \hat{f}(\theta')$, de loi normale, d'espérance m^* et de variance $\sigma_{m^*}^2$,

$$\frac{1}{2} L(\theta,\theta') = 1 - \Phi\left(\frac{\hat{f}(\theta') - \hat{m}^*}{\sigma_T(\theta,\theta')}\right) \qquad , \ \sigma_T(\theta,\theta') > 0$$

avec $\sigma_T^2(\theta,\theta') = \sigma_{m^*}^2 + \sigma_e^2(\theta') + \sigma_K^2 d(\theta',\theta)^{2\alpha}$

La fonction Φ désigne la fonction de répartition d'une loi normale centrée réduite. Nous prendrons la convention $L(\theta,\theta') = 1_{\hat{f}(\theta') = \hat{m}^*}$ dans le cas où $\sigma_T(\theta,\theta') = 0$, en l'absence de bruit et lorsque $\sigma_K = 0$ ou $\theta = \theta'$. Nous n'évoquerons pas en détail la détermination très classique, en un point exploré θ, de l'estimateur $\hat{f}(\theta)$ de $f(\theta)$, ni de l'estimateur de la variance de $\hat{f}(\theta)$.

L'usage de la moyenne et de la variance empirique non biaisée donnerait, à partir de n_0 observations de $F(\theta)$ notées $F_1(\theta),\ldots,F_{n_0}(\theta)$, $n_0 \geq 2$:

$$\hat{f}(\theta) = \frac{1}{n_0}\sum_{i=1}^{n_0} F_i(\theta)$$

$$\sigma_e^2(\theta) = \frac{1}{n_0}\hat{\sigma}_F^2(\theta), \text{ avec } \hat{\sigma}_F^2(\theta) = \frac{1}{n_0 - 1}\sum_{i=1}^{n_0}\left(F_i(\theta) - \hat{f}(\theta)\right)^2$$

$\sigma_F^2(\theta)$ désigne ici la variance de $F(\theta)$, et $\hat{\sigma}_F^2(\theta)$ un estimateur de cette variance. Des raffinements peuvent être envisagés afin de tenir compte de l'erreur d'estimation de σ_F. Notons surtout la nécessité d'un paramètre $n_0 \geq 2$, nombre de tirages requis pour l'estimation de la variance empirique $\sigma_e^2(\theta)$. Dans le cas déterministe, nous pouvons fixer $n_0 = 1$, $\hat{f}(\theta) = f(\theta)$ et $\sigma_e^2(\theta) = 0$. Mais dans le cas général, ce paramètre n_0 restera une entrée de l'algorithme. Eventuellement, l'usage d'une hypothèse de répartition gaussienne pour $F(\theta)$ peut permettre de ne pas mémoriser l'intégralité des tirages de $F(\theta)$, mais de simplement mémoriser le nombre de tirages précédents, leur somme et la somme de leurs carrés.

Pour l'estimation de m^* et de $\sigma_{m^*}^2$, nous retiendrons en première approche les valeurs obtenues pour $\hat{f}(\theta^*)$ et $\sigma_e^2(\theta^*)$ au dernier point optimal θ^* rencontré parmi les tirages réalisés itérativement par l'algorithme, bien que là encore des perfectionnements puissent être suggérés.

Remarquons que le potentiel peut être adapté à d'autres recherches que la recherche d'optimum. Ainsi, si l'on recherchait les paramètres θ tels que $f(\theta)$ appartienne à un ensemble $A \subset R$, on pourrait utiliser un potentiel proportionnel à $P\left[\tilde{f}_{\theta'}(\theta) \in A\right]$. L'algorithme peut ainsi être très facilement adapté à la recherche de racines.

Potentiel d'un point Considérons un point θ à l'intérieur d'une zone Z, et cherchons à définir une mesure d'autant plus grande qu'un minimum global de f peut être observé au point θ. Ce point θ est muni de $d + 1$ potentiels directionnels en direction des $d + 1$ sommets délimitant la zone Z, sommets notés ici θ_i, $i \in \{1,\ldots,d+1\}$. Ces potentiels marquent chacun la probabilité, selon la modélisation de l'observateur, que la fonction f franchisse le seuil m^* sur l'un des segments $[\theta, \theta_i]$. Les potentiels de points de différentes zones sont destinés à être comparés entre eux. La principale exigence est ici de respecter la logique booléenne selon laquelle si le minimum ne peut pas être atteint en direction d'un sommet, alors il est exclu qu'il soit atteint sur la zone. Une mesure commode respectant cette exigence est naturellement le produit, qui correspond bien au ET booléen lorsque les quantités $L(\theta, \theta_i)$ appartiennent à $\{0,1\}$.

Nous appellerons potentiel du point θ dans la zone Z la quantité :

$$\beta_Z(\theta) = \left(\prod_{i=1}^{d+1} L(\theta, \theta_i) \right)^{\gamma}$$

Nous avons choisi ici d'introduire une fonction lien $g(x) = x^{\gamma}$, pour $\gamma > 0$, qui préserve les propriétés requises de logique booléenne du produit, et permet de modifier le comportement de l'algorithme. Le choix de γ, régissant la convexité de g, permet de distordre les potentiels et de moduler ainsi l'importance relative accordée à la zone de plus grand potentiel. Au niveau de l'impact de la distorsion, pour un coefficient γ très grand, le tirage d'une zone reviendrait au tirage systématique de la zone dont le potentiel est le plus grand. Un coefficient γ faible conduirait à davantage explorer des zones de potentiel médian. Le facteur γ s'interprète comme un facteur de priorité attribuée aux zones de grand potentiel.

Nous avons constaté que l'effet de γ était modeste pour des valeurs raisonnables de paramètres (cf. sous-section II-4 de ce chapitre). Considérant par ailleurs une certaine redondance de ce paramètre avec les paramètres (σ_K, α) du potentiel, nous opterons sauf mention contraire pour :

$$\gamma = 1/(d+1)$$

qui a le mérite de clarifier l'interprétation du potentiel : $\beta_Z(\theta)$ correspondra simplement à la moyenne géométrique des potentiels directionnels. Le potentiel d'un point s'interprétera alors comme un potentiel directionnel moyen.

Le choix d'une mesure pour le potentiel d'un point justifierait de nombreuses études. S'agissant d'agrégation de potentiels directionnels et de fonctions liens, d'autres distorsions de probabilités facilement utilisables pourront être trouvées dans Bienvenüe et Rullière [2010]. Une autre piste pour cette agrégation est l'usage de copules, ou encore celle de la logique floue (cf. Zadeh [1965]). Nous étudions actuellement d'autres mesures de ce potentiel (cf. Rullière et al. [2010]), basés sur l'agrégation des variances directionnelles $\sigma_T(\theta, \theta')$, ainsi que la modélisation de $\tilde{f}$ par des champs aléatoires, en lien avec la théorie du Krigeage (cf. par exemple Krige [1951], Jones et al. [1998]). Ces mesures présentent l'avantage d'éliminer l'agrégation de potentiels directionnels, au prix d'un modèle d'une complexité parfois accrue.

Il faut néanmoins tempérer l'impact du choix d'une mesure de potentiel : les potentiels des différents points serviront à établir une hiérarchie des zones les plus susceptibles de contenir un optimum global, afin de choisir laquelle explorer en priorité. Le choix de la fonction lien g croissante modifiera les priorités d'exploration, mais ne modifiera pas la hiérarchie elle-même, et celle-ci dépendra en très grande partie de l'éloignement de f à la valeur de l'optimum.

Si les potentiels directionnels appartenaient à $\{0,1\}$, un potentiel permettrait simplement de dire si il est possible ou non que f possède un minimum global sur une zone, compte tenu des sommets adjacents et de la valeur estimée du minimum. Cela correspond à l'idée développée dans les algorithmes mettant en oeuvre une arithmétique d'intervalle (cf. Wolfe [1996]). Le potentiel proposé ici doit être vu comme une simple mesure, à valeur sur [0, 1], permettant

d'étendre une logique d'arithmétique d'intervalle qui conduirait à des potentiels définis sur $\{0,1\}$.

Potentiel d'une zone En pratique, une zone sera d'autant plus susceptible de contenir un minimum plus petit que la valeur estimée de m^* si sa surface est grande et si ses points ont un potentiel élevé. Une mesure logique de la "surface probable" d'une zone Z est donnée par :

$$\overline{\beta}(Z) = \int_{\theta \in Z} \beta_Z(\theta) d\theta$$

Bien que le calcul de cette intégrale puisse être approché par des techniques de simulation, nous avons préféré retenir comme mesure du potentiel d'une zone la mesure suivante :

$$\beta(Z) = V(Z).\beta_Z(\theta_{\beta_Z})$$

θ_{β_Z} isobarycentre des sommets de la zone Z,

où $V(Z)$ est le volume de la zone Z (hypervolume dans le cas $d > 3$).

Cette solution a le mérite de la simplicité, offre l'avantage d'être très rapide et de ne pas être aléatoire. Un calcul d'un $\beta(Z)$ précédemment mené n'aura donc pas à être réitéré si la zone Z n'est pas modifiée. En outre, l'objectif est de comparer les coefficients β entre eux : le calcul fin de l'intégrale a peu de raisons explicites de beaucoup perturber la hiérarchie entre les différentes zones. Enfin, l'isobarycentre nous a paru bien rendre compte de l'erreur de grille au sein de la zone, et facilitera l'interprétation future du potentiel d'une zone.

S'agissant du calcul du volume $V(Z)$, dans le cas où $d = 2$, les zones sont des triangles et un volume $V(Z)$ est donné par la formule de Héron. Dans le cas général, le déterminant de Cayley-Menger donne le volume exact de la zone (cf. Sommerville [1958]). Plus simplement, dans le cas d'une séparation d'une zone en deux volumes égaux, ce qui sera ici le cas, la simple mémorisation du volume de la zone à subdiviser permet de déduire immédiatement le demi-volume de chaque zone fille.

Choix d'une zone en fonction des potentiels Comme nous l'avons évoqué, le choix d'une zone Z^+ parmi un ensemble z de zones se fera de la façon suivante : la probabilité de piocher une zone sera proportionnelle au potentiel de chaque zone.

Notons $z = \{Z_1,...,Z_n\}$ l'ensemble des zones dans lequel doit être piochée Z^+. Si $\{U_v\}_{v=1,2,...}$ désigne une suite de variables aléatoires de loi uniforme sur [0,1], mutuellement indépendantes, piocher une zone parmi n avec une probabilité fixée au prorata de son potentiel revient à choisir, à une étape v de l'algorithme :

$$Z^+ = Z_{k^*(U_v)}$$

avec :

$$k^*(u) = min\{k \in \{1,..,n\}, B_k \geq u.B_n\}$$

et :

$$B_k = \sum_{i=1}^{k} \beta(Z_i),\ k \in \{1,...,n\}$$

Ce choix découle de plusieurs idées. D'une part, l'idée d'explorer de façon uniforme la zone de recherche lorsque les potentiels sont égaux, d'autre part l'idée de préserver, à la façon d'un

recuit simulé, la possibilité d'exploration de zones *a priori* peu prometteuses. Nous imaginons, en présence d'un très grand nombre de très petites zones, que cette solution limite le risque de confiner les points de tirage dans le voisinage d'un unique minimiseur, et favorise ainsi une certaine prudence dans l'exploration de la fonction. D'autres choix possibles sont évoqués dans la section d'applications numériques (sous-section II-4 de ce chapitre).

II-2-3 Choix des paramètres (σ_K , α) du potentiel

Au moyen du potentiel que nous avons défini, nous avons transféré une part de la subjectivité du choix du prochain point de tirage sur le choix de quelques paramètres, au premier rang desquels se trouvent les coefficients de type Hölder α et σ_K . Le choix ou l'estimation de ces paramètres est un problème délicat qui nécessiterait à lui seul une étude poussée, et que nous présentons ici de façon simplifiée.

Le choix des paramètres α et σ_K dépend de la connaissance de la fonction f considérée, ainsi que, lorsque celle-ci s'avère insuffisante, de la prudence de l'observateur. Le choix du paramètre α est un choix de modèle : envisageons-nous que f puisse varier brutalement sur un très petit intervalle comme le ferait une trajectoire de mouvement brownien, ou de façon plus régulière, comme une fonction lipschitzienne ?

Considérons deux points distincts θ et θ' , en ignorant dans un premier temps l'erreur d'échantillonnage aux sommets explorés. Si seul θ' est exploré, l'observateur suppose *a priori* que $Y_{\theta,\theta'} = \tilde{f}_{\theta'}(\theta) - f(\theta')$ est distribué selon une loi normale centrée, d'écart-type $\sigma_K d^{\alpha}$, où $d = d(\theta,\theta')$, d >0. Après exploration des deux points, l'observateur dispose d'une observation (d,y) de ce couple (d,Y) . En supposant que sont collectées *n* réalisations (d_i,y_i), en les supposant de surcroît mutuellement indépendantes, une estimation maximum de vraisemblance de σ_K en connaissance de α conduirait à :

$$\sigma_K^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i^2}{d_i^{2\alpha}}$$

et une estimation maximum de vraisemblance de α en connaissance de σ_K conduirait à α tel que :

$$\sum_{i=1}^{n}\frac{y_i^2\, ln(d_i)}{\sigma_K^2 d_i^{2\alpha}} = \sum_{i=1}^{n} ln(d_i)$$

Si l'on considère que, en tout point exploré θ , $\hat{f}(\theta)$ subit une erreur d'estimation, on peut observer $Y'_{\theta,\theta'} = \left(\tilde{f}_{\theta'}(\theta) + N_\theta\right) - \left(f(\theta') + N_{\theta'}\right)$, où N_θ et $N_{\theta'}$ sont des variables aléatoires indépendantes, de lois normales, de variances respectives σ_θ^2 et $\sigma_{\theta'}^2$. Alors $Y'_{\theta,\theta'}$ est supposée distribuée selon une loi normale centrée, de variance $\sigma_e^2 + \sigma_K^2 d^{2\alpha}$, où $\sigma_e^2 = \sigma_\theta^2 + \sigma_{\theta'}^2$ représente une erreur d'estimation, connue après exploration des deux points puisque mesurée aux points d'observation.

L'estimation maximum de vraisemblance de σ_K (sachant α) et de α (sachant σ_K), conduit respectivement à σ_K et α tels que :

$$\sum_{i=1}^{n} \frac{d_i^{2\alpha}}{\sigma_{e_i}^2 + \sigma_K^2 d_i^{2\alpha}} = \sum_{i=1}^{n} \frac{y_i^2 d_i^{2\alpha}}{\left(\sigma_{e_i}^2 + \sigma_K^2 d_i^{2\alpha}\right)^2}$$

$$\sum_{i=1}^{n} \frac{d_i^{2\alpha} ln(d_i) y_i^2}{\left(\sigma_{e_i}^2 + \sigma_K^2 d_i^{2\alpha}\right)^2} = \sum_{i=1}^{n} \frac{d_i^{2\alpha} ln(d_i)}{\sigma_{e_i}^2 + \sigma_K^2 d_i^{2\alpha}}$$

Dans la pratique, la répétition des étapes d'estimation de σ_K sachant α puis de α sachant σ_K a rapidement convergé dans tous les cas que nous avons testé (cf sous-section II-4 de ce chapitre).

Cette estimation nous a conduits à des résultats très proches de ceux obtenus en maximisant directement la log-vraisemblance de l'échantillon :

$$ln\, V(\sigma_k, \alpha) = -\frac{n}{2} ln(2\pi) - \sum_{i=1}^{n} \frac{1}{2} ln\left(\sigma_{e_i}^2 + \sigma_K^2 d_i^{2\alpha}\right) - \sum_{i=1}^{n} \frac{y_i^2}{2\left(\sigma_{e_i}^2 + \sigma_K^2 d_i^{2\alpha}\right)}$$

Au fur et à mesure des nouveaux tirages, chaque nouveau point θ_0 exploré conduit, en direction des d + 1 sommets de la zone, à d + 1 nouvelles réalisations de variables aléatoires de loi normale centrée, d'écart-type fonction de la distance, et il est alors possible de corriger à chaque étape une valeur *a priori* de α et de σ_K.

Il faut ici noter que rien n'interdit de faire varier les coefficients α et σ_K en fonction de la zone considérée, et l'estimation des coefficients pourrait se faire sur chaque zone Z en affectant chaque réalisation de la variable aléatoire $Y'_{\theta,\theta'}$ (sur d'autres zones) de poids d'autant plus élevés que la θ et θ' sont proches de la zone considérée Z. L'estimation de coefficients de Hölder à partir d'observations de tirages de F est un vaste sujet (cf. Blanke [2002] pour un article traitant de ce type d'estimation sur des processus). Les choix numériques concrets des paramètres de l'algorithme seront détaillés dans la sous-section d'application numérique II-4.

Enfin, le raisonnement tenu jusqu'à présent se basait sur le fait que pour deux points d'une zone, l'observateur représentait la variation des pentes entre ces deux points par une variable aléatoire dont l'écart-type était une fonction croissante de la distance. Même si cela n'est pas requis pour l'implémentation de l'algorithme, un élément susceptible de fournir a posteriori des informations sur l'hypothèse prise est le suivant : pour une zone Z, nous définissons une variable aléatoire IR_α :

$$IR_\alpha(Z) = \frac{f(\theta_1) - f(\theta_2)}{d(\theta_1, \theta_2)^\alpha}$$

θ_1, θ_2 vecteurs aléatoires distincts issus de tirages indépendants, uniformes sur Z.

Précisons ici les modalités des tirages. D'une part, le tirage uniforme d'un vecteur θ sur une zone Z peut s'opérer facilement en choisissant des coordonnées barycentriques de façon uniforme dans un simplexe unité : ainsi, si ω_j est la $j^{ème}$ coordonnée barycentrique de θ dans

la zone Z, $j \in \{1,...,d+1\}$, nous pourrons prendre $\omega_j = e_j / e_{tot}$, avec $e_{tot} = \sum_{j=1}^{d+1} e_j$, pour un ensemble $\{e_j\}_{j\in\{1,...,d+1\}}$ de variables aléatoires mutuellement indépendantes, de loi exponentielle de paramètre 1 (ce qui revient à prendre les intervalles successifs de statistiques d'ordre d'un vecteur uniforme). Le lecteur intéressé par ce type de tirages pourra consulter notamment Smith et Tromble [2004]. D'autre part, nous conviendrons que (θ_1, θ_2) est le premier couple distinct issu de tirages uniformes (du fait de la précision arithmétique finie des ordinateurs, la probabilité que les deux points soient confondus peut être non nulle en pratique, bien qu'extrêmement faible).

Selon le modèle présenté, à défaut d'autres informations, l'observateur suppose qu'il existe une quantité $\alpha \geq 0$ telle que la variation de f entre θ_1 et θ_2 ne s'explique que par la distance $d(\theta_1,\theta_2)^\alpha$, et pour lequel l'écart-type de $IR_\alpha(Z)$ serait fini (ce qui n'est pas acquis pour tout α bien sûr). Sous l'hypothèse prise, cet écart-type fournira une information sur les variations de pentes auxquelles l'utilisateur pourra s'attendre. En particulier, si $IR_\alpha(Z)$ est supposé distribué de façon gaussienne, alors son écart-type empirique correspondra à un estimateur de σ_K en connaissance de α. Des distributions de IR_α seront illustrées dans la sous-section II-4.

II-2-4 Choix du sommet de scission

Une fois une zone Z^+ choisie parmi un ensemble de zones z, il va s'agir de scinder la zone en ajoutant un point à l'intérieur de celle-ci. De par nos choix précédents, ce point se situera sur l'un des segments de l'enveloppe convexe de la zone.

Pour la zone Z^+, nous noterons $C = \{(\theta_{j_1},\theta_{j_2})\}_{j_1,j_2}$ l'ensemble des couples de sommets distincts de Z^+. Le sommet de séparation retenu peut être choisi de nombreuses façons.

La solution que nous retiendrons consiste ici à sélectionner, parmi l'ensemble des couples possibles, l'un de ceux qui maximisent la longueur du segment. Si $\theta_1,...,\theta_{d+1}$ désignent les sommets de la zone Z^+, nous choisirons comme sommet de séparation de la zone le sommet θ^+, isobarycentre du segment $(\theta_{j_1},\theta_{j_2})$ tel que :

$$(\theta_{j_1},\theta_{j_2}) = arg \max_{(\theta_i,\theta_j)\in C} d(\theta_i,\theta_j)$$
$$C = \{(\theta_i,\theta_j)\}_{i\in\{1,...,d\},\, j\in\{i+1,...,d+1\}}$$

Dans le cas de plusieurs segments de longueur maximale, nous conviendrons que *arg max* désignera un segment choisi de façon uniforme parmi l'ensemble (fini) des segments de longueur maximale.

Choisir le point de segmentation sur un segment de longueur maximale a le mérite de limiter l'apparition de simplexes très déséquilibrés dans leurs longueurs de segment : en effet, la connaissance de l'appartenance de l'optimum à une zone est moins utile si cette zone est très allongée dans certaines directions. Le choix d'un sommet conduisant également à la partition des zones adjacentes, l'apparition de simplexes très déséquilibrés, si elle est ainsi limitée,

n'est toutefois pas absolument exclue. Une idée de la forme des zones ainsi obtenues en dimension 2 est illustrée par la figure 47, qui a été construite par scission du segment de plus grande longueur de la zone choisie.

Lorsque des simulations sont opérées au point θ^+ du segment choisi de séparation, la scission des zones se fait ainsi :
– L'ensemble des zones z^+ contenant ce segment est déterminé (celles qui contiennent les deux sommets du segment à la fois).
– Chacune des zones Z de l'ensemble z^+ est scindée en deux zones Z_1 et Z_2 comme indiqué dans le paragraphe II-2-1 de ce chapitre.

II-2-5 Schéma de l'algorithme de scission systématique

Les données en entrée de l'algorithme sont les paramètres de régularité α et σ_K, le nombre de simulations n_0 à opérer en chaque point (nombre de tirages requis pour l'estimation de la variance empirique $\sigma_e^2(\theta)$), le nombre d'étapes n de l'algorithme. Nous supposons également que la zone de recherche Z_0 est connue et que les premiers tirages ont été réalisés aux sommets de cette zone.

L'algorithme 1 récapitule le procédé général de scission des zones. L'algorithme permet de construire à chaque étape j l'ensemble des zones z_j formant une partition de la zone de recherche initiale Z_0. Par construction, les nouveaux sommets de scission sont toujours explorés, de sorte que l'on a exploré, en fin d'algorithme, l'ensemble des sommets de la partition finale z_n. Nous rappelons ici que l'ensemble des sommets d'une zone Z est noté $S(Z)$. L'algorithme est ici présenté de façon synthétique. Chacune des étapes de l'algorithme est détaillée dans les sections précédentes. L'exploitation des nombreuses données disponibles en sortie de l'algorithme est détaillée dans la sous-section II-2-6 ci-après.

Algorithme 1 : algorithme à scission systématique
Entrée: σ_K, α, n_0, n
Entrée: $z_0 = \{Z_0\}$
pour $j = 0$ à $n-1$
choix d'une zone Z^+ *de* z_j (cf. § II-2-2)
calculer $\hat{m}^*$ et σ_{m^*}
$\forall Z_i \in z_j$, calculer $\beta(Z_i)$
piocher une zone Z^+ en fonction des $\{\beta(Z_i)\}_{Z_i \in z_j}$
choix d'un sommet de scission θ^+ *de* Z^+ (cf. § II-2-4)
piocher un sommet de scission θ^+
scinder les zones contenant θ^+
calculer z_{j+1}, ensemble des nouvelles zones
tirages au sommet $\theta^+ \in Z^+$
calculer n_0 tirages de $F(\theta^+)$
mise à jour facultative du couple (σ_K, α), éventuellement par zone (cf. § II-2-3)

fin pour

Sortie: $\hat{m}^*$, σ_{m^*}

Sortie: $\forall Z \in z_n$, $V(Z)$, $\beta(Z)$

Sortie: $\forall Z \in z_n$, $\forall \theta \in S(Z)$, $\hat{f}(\theta)$, $\sigma_e(\theta)$, $\beta_Z(\theta)$

II-2-6 Résultat final et critère de convergence

Zones de confiance Nous supposons ici que l'algorithme de scission systématique est terminé, par atteinte du critère d'arrêt proposé (nombre suffisant de tirages ici) : plus aucun tirage de F ne sera donc opéré.

Un estimateur de la valeur m^* de l'unique minimum global de f est fourni par l'algorithme. Pour autant, nous recherchons essentiellement à agir sur les paramètres, c'est-à-dire à obtenir l'ensemble des paramètres susceptibles de conduire à ce minimum, ce qui nous donnera également une indication sur la fiabilité du résultat obtenu et sur les investigations futures à opérer, par exemple pour départager deux candidats potentiels. Nous cherchons donc un ensemble discret (car il doit être traité numériquement) s'approchant (selon une mesure qui sera définie ultérieurement) de :

$$S_x = \{\theta \in \Theta, E[F(\theta)] \leq x\}$$

pour tout x appartenant à un voisinage de m^*.

A l'issue de l'algorithme, le domaine de recherche initial Z_0 est scindé en un ensemble de zones z. La définition des zones de confiance sera facilitée si nous définissons le potentiel d'un point y compris sur les frontières qui peuvent appartenir à plusieurs zones à la fois (sur les faces des simplexes). Les points appartenant à plus d'une zone formeront un ensemble de volume nul. Toutefois, afin de lever toute ambiguïté pour ces points particuliers, nous définirons pour chaque point une unique valeur de potentiel :

$$\beta_z(\theta) = \max_{Z \in z, \theta \in Z} \beta_Z(\theta)$$

L'ensemble des paramètres admissibles, que nous nommerons également zone de confiance, sera défini pour un seuil $s \in [0,1]$ comme l'ensemble des points candidats θ pour lesquels $f(\theta)$ est potentiellement inférieur à m^* :

$$\hat{S}_{m^*,s} = \{\theta \in \Theta_0, \beta_z(\theta) \geq s\}$$

L'ensemble Θ_0 des points candidats pourra être l'ensemble des points pour lesquels des tirages ont été opérés, ou bien, lorsque cela facilite l'usage futur de $\hat{S}_{m^*,s}$, un ensemble de points régulièrement espacés, pour un pas donné strictement positif δ :

$$\Theta_0 = \{\theta \in \Theta, \frac{1}{\delta}\theta \in \mathrm{N}^d\}$$

Lorsque s est égal à 0, tous les points de Θ_0 sont retenus. Lorsque s augmente, l'ensemble $\hat{S}_{m^*,s}$ se restreint à l'ensemble des paramètres pour lesquels $\hat{f}(\theta)$ est très proche de m^*.

En résumé, nous obtenons finalement :
– la valeur $m*$ de l'optimum de f ,
– l'ensemble des paramètres $\hat{S}_{m*,s}$ susceptibles de conduire à cet optimum.

Critère de convergence en connaissance du résultat Nous allons dans un premier temps chercher un critère de convergence de l'algorithme lorsque nous connaissons l'ensemble de solutions S_{m*}. Un tel critère facilitera la compréhension de l'algorithme sur des fonctions de test, et constituera un outil de comparaison de différents algorithmes.

Nous cherchons à produire un ensemble $\hat{S}_{m*}$ qui donne une représentation fidèle de S_{m*}. Par fidèle, nous imaginons d'une part que tout point du véritable ensemble solution S_{m*} doit être proche d'un point de l'ensemble solution proposé $\hat{S}_{m*}$: $\hat{S}_{m*}$ doit être suffisamment grand (condition n° 1). D'autre part, l'ensemble proposé ne doit pas non plus contenir de points trop éloignés des véritables solutions, et tout point de l'ensemble proposé $\hat{S}_{m*}$ doit être proche d'un point de S_m : $\hat{S}_{m*}$ ne doit pas être trop grand (condition n° 2). Cela nous conduira à proposer, comme distance entre les deux ensembles S_{m*} et $\hat{S}_{m*}$ la distance de Hausdorff, pour $X \subset \Theta$, $Y \subset \Theta$:

$$d^H(X,Y) = max\{ \sup_{y\in Y} \inf_{x\in X} d(x,y), \sup_{x\in X} \inf_{y\in Y} d(x,y)\}$$

Si $X = \hat{S}_{m*,s}$ est l'ensemble proposé et $Y = S_{m*}$ est l'ensemble cible, alors majorer le premier terme du max indique que tout point de la cible Y est proche d'un point de X (condition n° 1). Majorer le second terme du max indique que tout point de X est proche d'un point de Y (condition n° 2). Finalement, un point que nous proposons comme solution ne doit pas être trop éloigné d'une solution réelle, et une solution réelle ne doit pas être trop éloignée d'un point proposé comme solution. La distance de Hausdorff est donc parfaitement adaptée à l'objectif recherché. Cette distance représente, pour le pire point de l'un des ensembles, la distance de ce point à l'autre des deux ensembles : elle fournit directement une idée de l'incertitude sur l'ensemble des paramètres conduisant à l'optimum. Nous obtenons donc un critère de convergence en connaissance du résultat recherché S_{m*} :

$$\rho_1(s) = d^H(\hat{S}_{m*,s}, S_{m*})$$

Le critère précédant dépendant de la mesure choisie pour le potentiel, nous proposons également l'usage d'une distance de Hausdorff partielle :

$$\rho_2 = \sup_{y\in S_{m*}} \inf_{x\in \Theta_e} d(x,y)$$

Où Θ_e représente l'ensemble des points explorés. Ce critère fournit la pire distance d'un point solution au plus proche point exploré. Il indique donc si tous les points solutions ont bien été explorés, et serait naturellement très bon si Θ_e recouvrait Θ. Ce critère n'a de sens que dans la mesure où le nombre de tirages total de F est limité : il pourra notamment servir à comparer différents algorithmes pour un même budget de tirages. Son avantage est de ne requérir que l'ensemble des points de tirage successifs de F. Une limite de ce critère est qu'un algorithme peut converger très vite vers un optimum global sans avoir exploré les zones candidates, et donc en ayant pris un risque important : trouver rapidement le vrai optimum

global d'une fonction n'indique pas si la méthode est prudente ou non, et seule la considération de la variabilité de la fonction entre les points de tirage permet de trancher cette question (en un sens l'usage d'un potentiel). Une autre limite de ce critère est qu'il ne tient pas compte de la précision de l'estimation de f aux points explorés. Nous avons néanmoins tenu à le présenter dans la mesure où il s'agit d'un des critères permettant de comparer l'algorithme proposé à d'autres algorithmes ne permettant pas le calcul de potentiels.

Il faut noter que du fait du caractère aléatoire de la fonction observée et de l'algorithme proposé, ces critères devraient varier selon les exécutions de l'algorithme. En toute rigueur, pour un critère ρ, la comparaison de plusieurs algorithmes sur une fonction test, pour un même budget de tirages autorisé, devrait requérir l'obtention des distributions de ρ pour chaque algorithme, puis l'usage d'un indicateur de risque, comme un quantile de ρ. Une piste pour le choix opérationnel des paramètres de l'algorithme est l'étude approfondie du comportement du critère choisi en fonction de ces paramètres.

Critère de convergence hors connaissance du résultat Dans la pratique, nous ne connaissons pas l'ensemble cible S_{m^*}, puisque nous cherchons précisément à en cerner les contours.

Si nous cherchons à assurer que la fonction est suffisamment bien connue sur chaque zone, un indicateur de convergence est simplement la quantité :

$$\rho_3 = \max_{Z \in z} \beta(Z)$$

avec le choix ici opéré $\beta(Z) = V(Z).\beta_Z(\theta_{B_Z})$, θ_{B_Z} isobarycentre de la zone Z (cf. section II-2 de ce chapitre).

Majorer cet indicateur garantira en effet que pour toute zone $Z \in z$:
– Soit que la présence d'un minimum global au point central est très peu probable, $\beta_Z(\theta_{B_Z})$ étant suffisamment faible.
– Soit que la zone est suffisamment petite et que le minimum potentiel a bien été exploré, $V(Z)$ étant suffisamment faible.

Enfin, si nous cherchons à assurer que chaque point ne conduira pas à un minimum plus petit que celui connu, pour un écart η donné, un critère pourra être :

$$\rho_4 = \max_{\theta \in \Theta_0} \beta_z^{(m^* - \eta)}(\theta)$$

Nous notons ici $\beta_z^{(m^*)}(\theta)$ le potentiel au point θ obtenu pour un minimum global estimé m^* (l'indice supérieur $^{(m^*)}$ était resté implicite jusqu'à présent pour ne pas alourdir les notations). D'autres critères peuvent naturellement être envisagés. Un élément incontournable dans la définition d'un critère est la modélisation de $f(\theta)$ ou de l'incertitude de $\hat{f}(\theta)$ en dehors des points θ d'observation : si f *est* ou *est supposée* extrêmement erratique, toute zone sera susceptible de contenir un point conduisant à un optimum global, et le critère devra être très différent de celui obtenu en supposant f très régulière. C'est un avantage des critères proposés ici, à l'exception de ρ_2, mais c'est également une limitation puisque ces critères ne s'appliqueront qu'aux algorithmes proposant une partition de l'espace en zones affectées d'un potentiel.

II-3 Une grille à pas variable avec retirage possible

Jusqu'à présent, chaque zone était systématiquement subdivisée en deux. Or, il peut être plus avantageux, plutôt que de subdiviser une zone, d'ajouter des simulations en des points déjà explorés. Ce sera l'objet de cette sous-section.

II-3-1 Critères de scission

Supposons que nous estimons une espérance et un intervalle de confiance de la fonction $f(\theta)$, aux points explorés θ_1 et θ_2. Supposons que θ_c soit le barycentre équipondéré (isobarycentre) entre ces deux points.

La question que l'on se pose est la suivante : au vu des intervalles de confiance de $f(\theta)$, connus aux points θ_1 et θ_2, et connaissant la variabilité de l'évolution de $f(\theta)$ sur ce segment, mesurée par les paramètres ($\alpha ; \sigma_K$), nous cherchons un critère permettant de déterminer la décision à prendre entre les suivantes :

– diminuer l'incertitude sur $\hat{f}(\theta_i)$ par de nouveaux tirages de F en θ_i, $i \in \{1,2\}$.
– scinder le segment en deux parties, en opérant des tirages de F au nouveau point θc.

Plusieurs critères peuvent être retenus pour opérer ce choix. Le choix du critère retenu au final dépend essentiellement de l'objectif poursuivi : s'assurer que le minimum ne peut être présent dans une zone sous l'hypothèse que la fonction objectif ne varie pas trop brusquement, minimiser le maximum des potentiels d'une zone, etc. Ce choix pouvant dépendre du problème considéré, nous détaillerons ci-dessous deux critères dont nous étudierons la performance dans la sous-section II-4 de ce chapitre.

Comparaison des erreurs de grille et d'estimation Considérons une zone Z. En un point θ de Z, le potentiel est la mesure retenue pour quantifier la probabilité que $\tilde{f}(\theta)$ soit plus petite que l'optimum global $\hat{m}*$ estimé. Cette mesure dépend à la fois de l'erreur d'estimation aux sommets explorés adjacents, et à la fois de l'erreur de grille, liée à l'éloignement entre θ et les sommets adjacents (et dépendant de la régularité supposée de $\tilde{f}$, mesurée par α et σ_K). Par souci de simplicité, l'incertitude liée à l'estimation du minimum global $m*$ ne sera pas ici scindée en erreur d'estimation et erreur de grille.

Considérons l'*erreur d'échantillonnage*. Cette erreur est liée à la mauvaise connaissance de la fonction f aux points explorés, du fait du nombre réduit de tirages et du bruit frappant f. Elle dépend essentiellement de la probabilité de présence du minimum aux sommets θ_i de Z du fait de trop larges intervalles de confiance aux points explorés. En l'absence de doute lié à l'évolution de la fonction entre ces points, cette probabilité peut être mesurée par le potentiel $\beta(\theta)$ calculé en l'absence d'erreur de grille :

$$\beta_e(\theta) = \beta(\theta)/_{\sigma_K=0}$$

Considérons l'*erreur de grille*. Cette erreur dépend essentiellement de la possibilité que la fonction évolue brusquement entre un sommet de Z et le point θ considéré. En l'absence de doute lié à la mauvaise connaissance de la fonction aux extrémités du segment, compte tenu de la seule mauvaise exploration de la fonction, cette erreur peut être estimée au point θ par le

potentiel en θ, calculé en l'absence d'erreur d'estimation aux d + 1 sommets θ_i de la zone Z :

$$\beta_g(\theta) = \beta(\theta) /_{\sigma_{e_i}=0, i=1,\ldots,d+1}$$

Envisageons de scinder un segment autour d'un point barycentre non encore exploré θ_c. Si le potentiel en ce point est grand essentiellement à cause de l'erreur de grille, il apparaîtra raisonnable de scinder le segment et d'explorer θ_c. Si le potentiel est grand essentiellement à cause de l'erreur d'estimation, il paraîtra plus raisonnable au contraire de réduire cette dernière en explorant davantage les sommets du segment. Le critère de scission que nous retiendrons sera donc le suivant :

Critère n° 1 : Scission si $\beta_g(\theta_c) \geq \beta_e(\theta_c)$

En dimension d = 1, si θ_c est l'isobarycentre de θ_1 et θ_2, alors nous pouvons noter $d_0 = d(\theta_c, \theta_1) = d(\theta_c, \theta_2)$. Si nous supposons $\sigma_e = \sigma_e(\theta_1) = \sigma_e(\theta_2)$, compte tenu de la définition retenue ici pour le potentiel d'une zone, nous montrons facilement que l'équivalence suivante est obtenue :

$$(\beta_g(\theta_c) \geq \beta_e(\theta_c)) \Leftrightarrow (\sigma_e \leq \sigma_K d_0^{\alpha})$$

Ce dernier critère s'interprète alors très aisément : la scission est opérée si l'erreur de grille $(\sigma_g = \sigma_K d_0^{\alpha})$ est supérieure à l'erreur d'estimation (σ_e). En dimension d > 1, le critère $\beta_g(\theta) \geq \beta_e(\theta)$ permet de tenir compte en un point θ des différences éventuelles de distances aux sommets adjacents, et des différences d'erreurs d'estimation aux différents sommets de la zone Z : Ce critère revient à opérer une moyenne particulière entre les erreurs d'estimation σ_e et les erreurs de grille σ_g en direction des différents sommets.

En toute logique, si F est aléatoire et si σ_K = 0, alors la supposition de la connaissance parfaite de f entre les points de tirage conduit à seulement mieux estimer $\hat{f}$ aux sommets de la zone Z. Par ailleurs, en l'absence d'erreur d'estimation (par exemple si F est déterministe), le critère conduit bien à une scission systématique, il devient naturellement inutile d'effectuer un retirage.

Meilleur potentiel après ajout de n_0 simulations Nous envisageons ici le cas où la répartition se fait en fonction de l'amélioration du potentiel maximal de la zone, scindée ou non. Nous supposons que n_0 simulations seront ajoutés sur un sommet de la zone choisie Z^+. Les erreurs d'estimation correspondent à des écarts-type de variables aléatoires supposées de loi normales. Il est donc simple d'estimer quelle sera la réduction de la variance empirique en cas d'ajout de n_0 simulations, avant même la réalisation de ces simulations. Ainsi, en un point déjà exploré θ, si n_θ tirages de F ont déjà eu lieu et si $\sigma_F(\theta)$ représente la variance empirique de F en ce point, nous pourrons proposer $\sigma_e(\theta) = \sigma_F(\theta) / \sqrt{n_\theta + n_0}$.

– Si nous envisageons d'ajouter n_0 simulations en un sommet existant θ de la zone choisie, il est donc aisé d'estimer la nouvelle erreur d'échantillonnage $\sigma_e(\theta)$ en cas d'ajout.

– Si nous envisageons d'ajouter n_0 points sur le barycentre θ_c du segment choisi, nous pouvons alors estimer $\sigma_F(\theta_c)$ par interpolation linéaire, avant l'exploration du point θ_c, et en déduire une estimation de l'erreur d'échantillonnage $\sigma_e(\theta_c) = \sigma_F(\theta) / \sqrt{n_0}$ en cas d'exploration.

Quel que soit le sommet choisi pour l'ajout parmi les d +1 sommets existants, nous pouvons donc donner un nouvel estimateur $\hat{\beta}(Z^+)$ du potentiel de la zone choisie après ajout de n_0 simulations. Si un sommet existant doit être privilégié pour l'ajout, il semble logique de convenir d'un ajout sur celui de ces d + 1 points qui conduit à minimiser cette valeur estimée, ce minimum étant noté $\hat{\beta}_{min}$:

$$\hat{\beta}_{min} = \min_{\theta_i \in S(Z^+)} \left\{ \hat{\beta}(Z^+) \Big|_{n_0 ajouts\ en\ \theta_i} \right\}$$

Si au contraire une scission est envisagée, deux zones Z_1 et Z_2 seront formées. La scission pourra être privilégiée, par exemple, si les potentiels estimés sur ces deux zones sont tous deux inférieurs à β_{min} :

Critère n° 2 : Scission si $max\left(\hat{\beta}(Z_1), \hat{\beta}(Z_2)\right) < \hat{\beta}_{min}$

L'esprit étant alors qu'une scission ne doit pas conduire à augmenter le maximum des $\beta(Z_i)$ sur l'ensemble des zones $\{Z_i\}$ à l'étape considérée : l'ajout se porte ainsi sur le sommet dont nous estimons qu'il minimise le maximum des potentiels sur l'ensemble des zones.

Cette répartition est également pleinement cohérente avec la définition choisie des coefficients β, et elle conduit, par construction, à une diminution du potentiel maximum estimé sur l'ensemble des zones, ce qui est de nature à faciliter les futures démonstrations de la convergence de l'algorithme, notamment si le critère de convergence ρ_3 est utilisé.

II-3-2 Schéma de l'algorithme à retirage possible

Algorithme 2 : algorithme à retirage possible

Entrée: σ_K , α , n_0 , n

Entrée: $z_0 = \{Z_0\}$

pour $j = 0$ à n -1

choix d'une zone Z^+ *de* z_j (cf. § II-2-2)

calculer $\hat{m}^*$ et σ_{m^*}

$\forall Z_i \in z_j$, calculer $\beta(Z_i)$

piocher une zone Z^+ en fonction des $\{\beta(Z_i)\}_{Z_i \in z_j}$

si critère scission (Z^+) vrai (cf. § II-3-1) *alors*

choix d'un sommet de scission θ^+ *de* Z^+ (cf. § II-2-4)

piocher un sommet de scission θ^+

scinder les zones contenant θ^+

calculer z_{j+1}, ensemble des nouvelles zones

sinon

choix d'un sommet de retirage θ^+ *de* Z^+ (cf. § II-3-1)

piocher un sommet déjà exploré $\theta^+ \in S(Z^+)$

$z_{j+1} = z_j$, les zones sont inchangées

fin si

tirages au sommet $\theta^+ \in Z^+$

calculer n_0 tirages de $F(\theta^+)$

mise à jour facultative du couple (σ_K , α), éventuellement par zone (cf. § II-2-3)

fin pour

Sortie: $\hat{m}*$ et σ_{m^*}

Sortie: $\forall Z \in z_n$, $V(Z), \beta(Z)$

Sortie: $\forall Z \in z_n, \forall \theta \in S(Z)$, $\hat{f}(\theta), \sigma_e(\theta), \beta_Z(\theta)$

L'algorithme 2 récapitule le procédé général de scission des zones. L'algorithme permet de construire à chaque étape j l'ensemble des zones z_j formant une partition de la zone de recherche initiale Z_0. Cet algorithme diffère de l'algorithme 1 par sa faculté d'opérer un nouveau tirage en un sommet déjà exploré de la zone à explorer Z^+, plutôt que de systématiquement scinder Z^+.

II-4 Applications numériques

II-4-1 Fonction test utilisée

Nous présenterons ici une application partant d'une fonction test connue, afin notamment de vérifier le bon comportement de l'algorithme. Par souci de lisibilité, cette application sera tout d'abord présentée en dimension $d = 2$ (mais l'algorithme présenté dans ce travail est proposé pour toute dimension $d \in \mathrm{N}^*$, et le problème de la montée en dimension est évoqué avec plus de détail dans Rullière et al. [2010]).

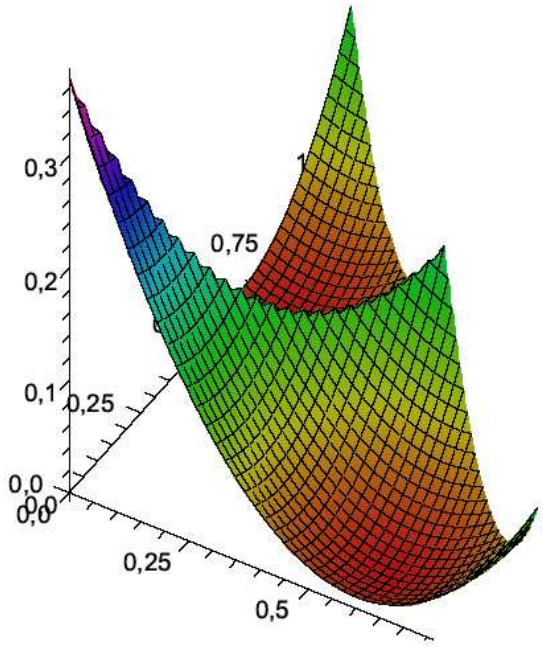


Fig. 48 : *Allure générale de la fonction test* $f(\theta)$ *pour* $\theta \in Z_0$

Nous considérerons ici la fonction suivante dans le cas $d = 2$ (correspondant par exemple à 3 poids d'allocation d'actifs se sommant à 1).

$$f(\theta) = \left(min(x,y) - 0.1\right)^2 + \left(max(x,y) - 0.6\right)^2$$
$$F(\theta) = f(\theta) + \sigma_B(U - 0.5)$$
$$\theta = (x,y)$$

avec U une variable aléatoire uniforme sur [0,1]. L'espérance f de la fonction F admet deux minima, l'un en $\theta_1^* = (0.1,0.6)$, l'autre en $\theta_2^* = (0.6,0.1)$, la valeur de f étant alors 0. A titre indicatif, afin d'imaginer les variations possibles de cette fonction, la fonction f atteint son maximum sur Z_0 en $\theta = (0,0)$ et nous avons alors $f(\theta) = 0.37$ (mais nous cherchons ici le minimum, non le maximum). Un bruit pouvant conduire à des variations d'amplitude de 0.1 entre deux points proches est donc assez élevé au regard du domaine de variation [0, 0.37] de la fonction sur la zone initiale de recherche. Un aperçu de la fonction f est donné dans le graphique 48.

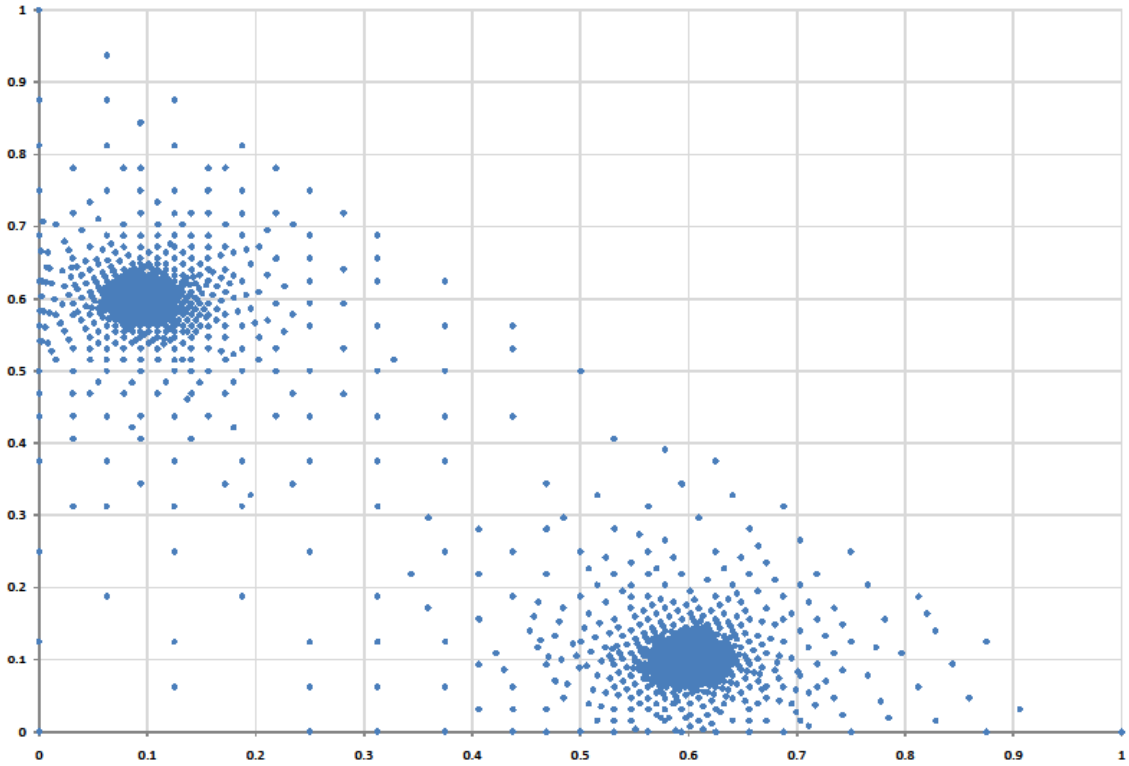


Fig. 49 : *Ensemble des points de simulations de* F *pour une variabilité* σ_K *= 2 et un niveau de bruit* σ_B *= 0 (scission systématique)*

Dans le cadre de l'algorithme à scission systématique, nous utiliserons ici toujours les paramètres suivants :
– le nombre n de tirages réalisés, sera n = 2000,
– le nombre de tirages en chaque point sera n_0 = 10,
– le coefficient de priorité sera $\gamma = 1/(d+1) = 1/3$, conduisant à une moyenne géométrique des potentiels directionnels pour le potentiel d'un point,
– le seuil η du critère de convergence ρ_4 sera η = 0.01.

Par ailleurs, pour chaque illustration, nous préciserons
– le niveau du bruit σ_B frappant la fonction f,
– les coefficients de variabilité σ_K et α,
– le seuil s du critère de convergence ρ_1.

A titre de remarque, le graphique 47 présentée précédemment a été obtenue avec la fonction F évoquée ci-dessus, à partir des paramètres (σ_K, α) = (1, 1.5) et σ_B = 0.1 (et γ = 1 dans ce seul cas).

II-4-2 Comportement de l'algorithme à scission systématique

Dans cette sous-section, nous allons observer le comportement de l'algorithme n° 1 à scission systématique, lorsque le niveau de variabilité σ_K varie, pour différents niveaux de bruit σ_B.

Dans ce paragraphe, le paramètre α est ici fixé, égal à 1.5, et correspond à un modèle où les pentes sont supposées suivre un mouvement brownien depuis un point d'accroche (cf. définition du potentiel directionnel dans les paragraphes II-2-2 et II-2-3 de ce chapitre).

A titre indicatif, nous mentionnerons dans la légende de certains graphiques la valeur obtenue pour le vecteur critère de convergence, $\vec{\rho} = (\rho_1(s), \rho_2, \rho_3, \rho_4(\eta))$, avec s égal à 10 % du potentiel maximal observé, et η = 0.01. Le détail du calcul de ces critères ainsi qu'une analyse plus détaillée des valeurs numériques obtenues pour ceux-ci seront abordés dans le paragraphe II-4-5 de ce chapitre.

Pour un niveau de variabilité supposée de la fonction σ_K = 2, illustré dans les graphiques 49, 50, et 51, lorsque le niveau de bruit passe respectivement par σ_B = 0, σ_B = 0.1 et σ_B = 0.3, nous observons bien l'exploration préférentielle des zones à proximité des minimiseurs de f. Bien entendu, cette exploration est plus étendue lorsque le niveau de bruit augmente, puisque du fait de ce bruit, l'assurance de l'absence d'autres optima locaux nécessite l'exploration d'une zone plus vaste. Il faut noter que dans la pratique, ce niveau de bruit de F est généralement une donnée exogène sur laquelle il n'est pas possible d'agir.

Pour un niveau de bruit fixé à σ_B = 0.1, illustré dans les graphiques 50, 52 et 53, lorsque la variabilité supposée de la fonction passe successivement par σ_K = 2, σ_K = 20 et σ_K = 100, nous observons un résultat parfaitement logique : si le comportement attendu de la fonction entre les points de tirage est supposé peu variable, l'algorithme se concentre sur les zones proches des minimiseurs supposés. Dans le cas inverse, une variabilité extrême σ_K = 100 conduit à une répartition quasi uniforme des points d'exploration : avec une telle variabilité, l'optimum global de la fonction peut en effet se trouver dans n'importe quelle zone. Il faut toutefois remarquer que la fixation d'un seuil de variabilité σ_K trop faible conduirait par définition à supposer que le comportement de la fonction est globalement connu entre les points explorés, ce qui conduit à négliger des zones pourtant susceptibles de contenir un optimum. L'arbitrage entre la rapidité de convergence vers les optima globaux et le risque d'avoir négligé une zone dépendra donc largement de ce paramètre σ_K.

Pour α fixé, la fixation du paramètre σ_K est donc un problème délicat, que nous aborderons dans le paragraphe II-4-4. Elle pourra se baser en particulier sur les explorations précédentes de la fonction, sur d'autres connaissances de celles-ci, ainsi que sur les contraintes d'implémentation et de niveau de risque accepté. Il faut noter à ce sujet que rien dans le modèle n'oblige à fixer cette valeur constante sur l'ensemble de la zone de recherche, et qu'il est également possible de tenir compte d'une variabilité particulière de la fonction sur certaines zones.

En résumé, pour l'algorithme n° 1 à scission systématique, l'algorithme vise essentiellement à répartir des points d'exploration, en délaissant temporairement certaines zones non susceptibles de contenir le minimum. Selon les paramètres choisis et la fonction optimisée, le comportement attendu est le suivant, conforme à ce que nous avons pu observer :

– Lorsque le bruit est important devant les variations de la fonction, ou lorsque l'incertitude sur la variabilité de la fonction σ_K est élevée, l'algorithme explore assez

uniformément la fonction, de façon similaire aux traditionnelles grilles de recherche à pas fixe évoquées précédemment.

– Lorsque le bruit n'est pas trop élevé et lorsque la variabilité de la fonction σ_K est faible, les points de tirage se concentrent sur les zones contenant les points solutions supposés, le comportement de la fonction étant supposé sans surprise entre les points déjà explorés.

Le choix du coefficient σ_K de variabilité de la fonction f dépend donc du but poursuivi : trop élevé, l'exploration de la fonction sera très poussée, et le temps de calcul pourra être élevé si nous visons une bonne connaissance d'un optimum de la fonction. Pour un σ_K trop faible, l'algorithme se focalisera très vite sur un optimum local, donc permettra un gain en terme de temps de calcul, à précision égale, mais le risque de ne pas explorer une zone susceptible de contenir un autre optimum, éventuellement meilleur, sera plus élevé.

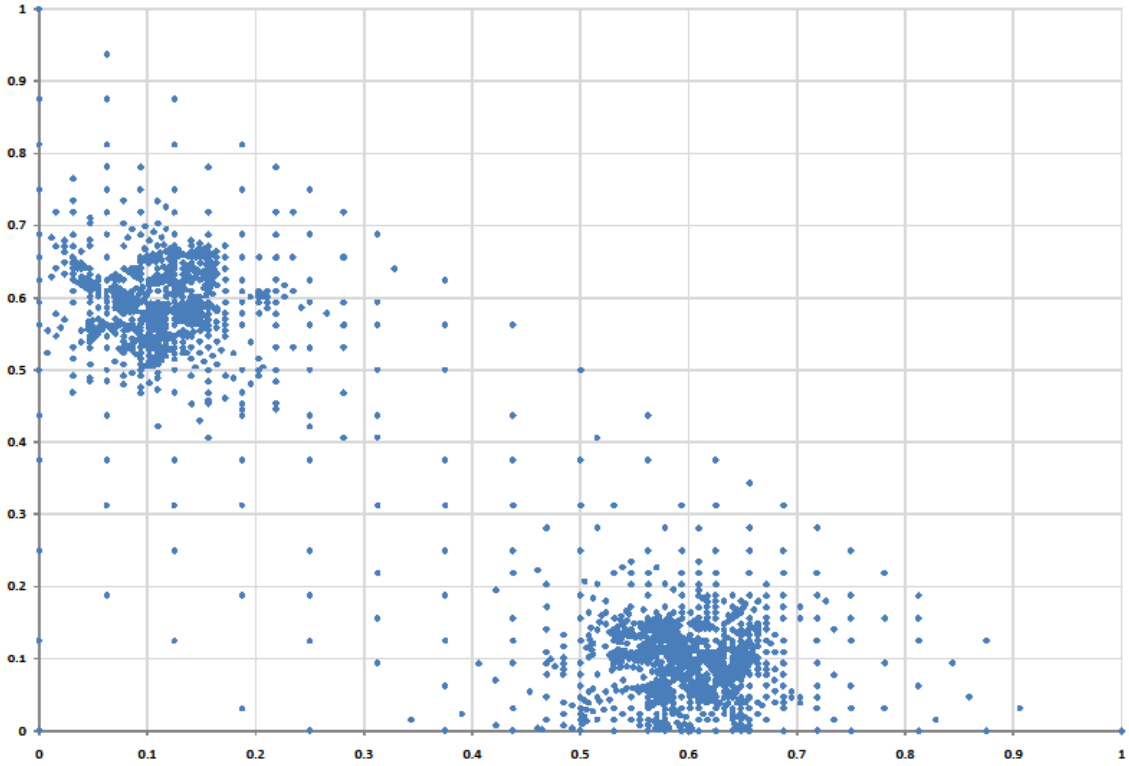


Fig. 50 : *Ensemble des points de simulations de* F *pour une variabilité* σ_K *= 2 et un niveau de bruit* σ_B *= 0.1 (scission systématique),* $\vec{\rho}$ *= (0.10, 3.5E-03, 6.8E-07, 1.9E-2)*

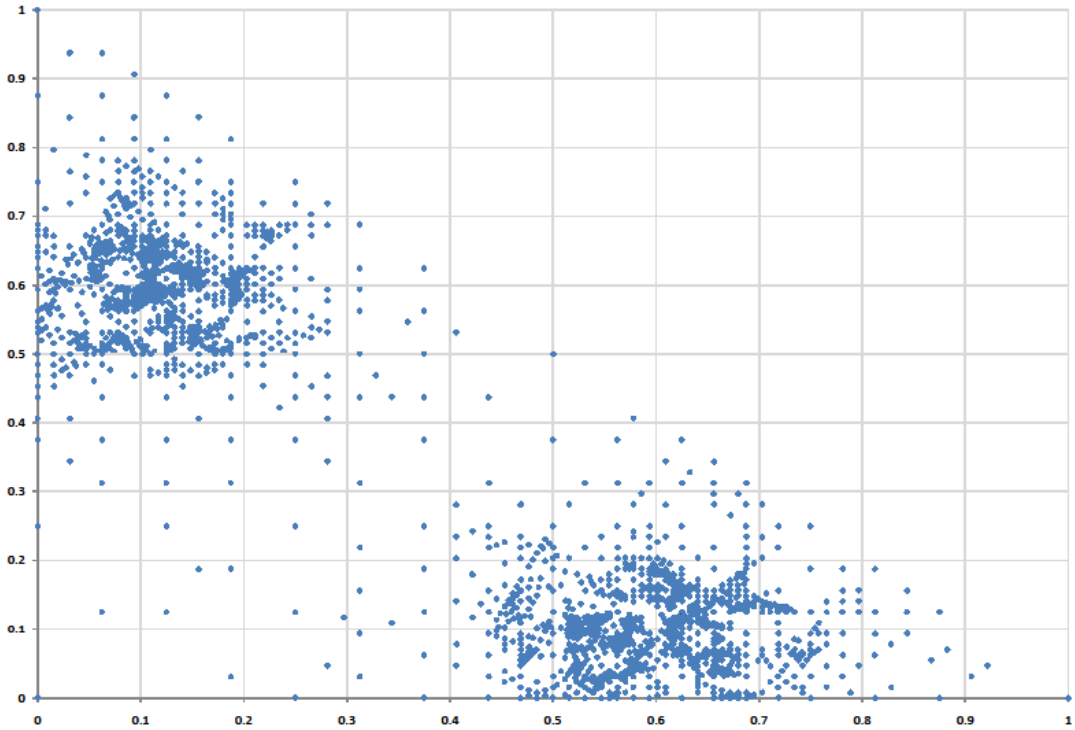


Fig. 51 : *Ensemble des points de simulations de* F *pour une variabilité* σ_K *= 2 et un niveau de bruit* σ_B *= 0.3 (scission systématique),* $\vec{\rho}$ *= (0.13, 5.9E-03, 1.6E-06, 5.8E-2)*

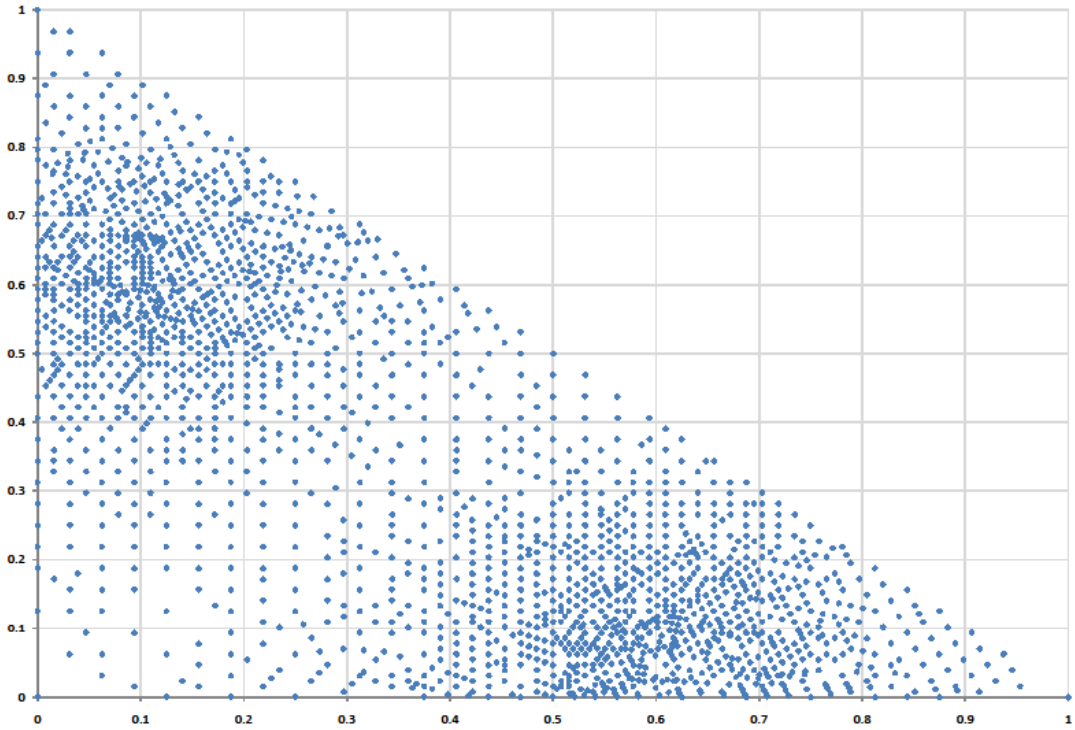


Fig. 52 : *Ensemble des points de simulations de* F *pour une variabilité* σ_K *= 20 et un niveau de bruit* σ_B *= 0.1 (scission systématique),* $\vec{\rho}$ *= (0.45, 2.2E-03, 7.3E-05, 0.72)*

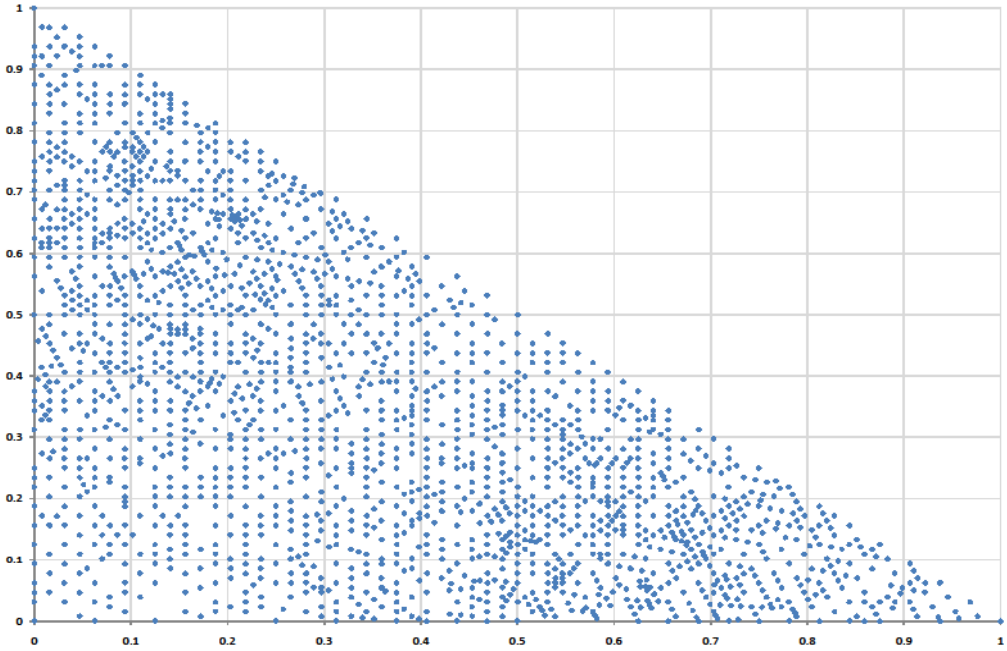


Fig. 53 : *Ensemble des points de simulations de F pour une variabilité σ_K = 100 et un niveau de bruit σ_B = 0.1 (scission systématique), $\vec{\rho}$ = (0.60, 8.8E-03, 5.9E-04, 0.97). La recherche de l'optimum d'une fonction constante bruitée conduit à un nuage de points d'allure proche de celui-ci.*

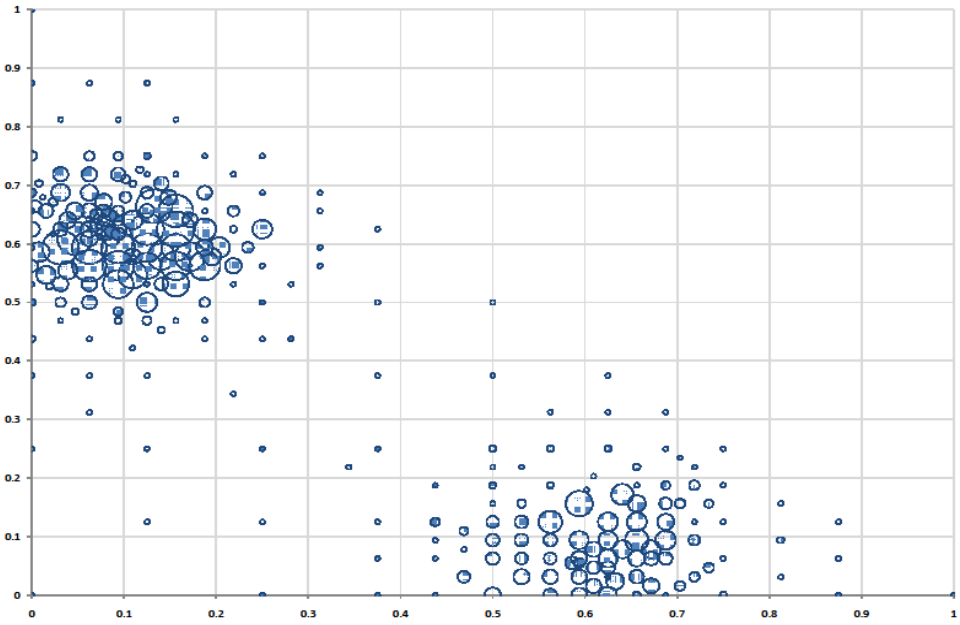


Fig. 54 : *Ensemble des points de simulation de F pour un niveau de bruit σ_B = 0.1, critère de scission n° 2, pour une variabilité (σ_K ; α) = (1, 1.5), et un nombre minimal de tirage n_0 = 10. La surface des bulles est proportionnelle au nombre de tirages de F en chaque point. $\vec{\rho}$ = (0.085, 8.8E-03, 3.1E-05, 8.2E-04)*

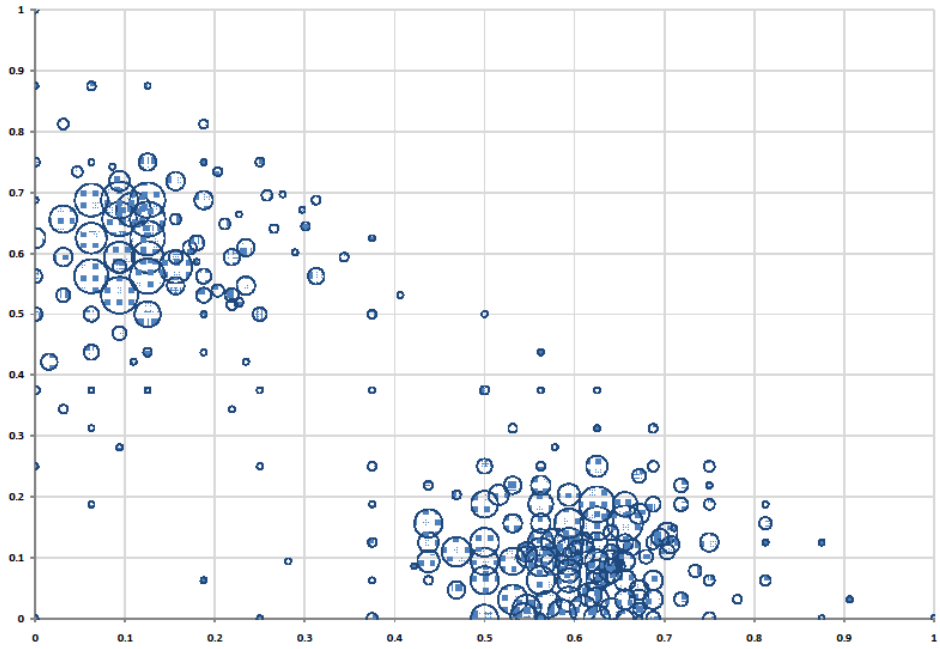


Fig. 55 : *Ensemble des points de simulation de* F *pour un niveau de bruit* σ_B *= 0.2, critère de scission n° 2, pour une variabilité (*σ_K *,* α *) = (1, 1.5), et un nombre minimal de tirage* n_0 *= 10. La surface des bulles est proportionnelle au nombre de tirages de* F *en chaque point.* $\vec{\rho}$ *= (0.10, 8.8E-03, 6.7E-05, 6.2E-03)*

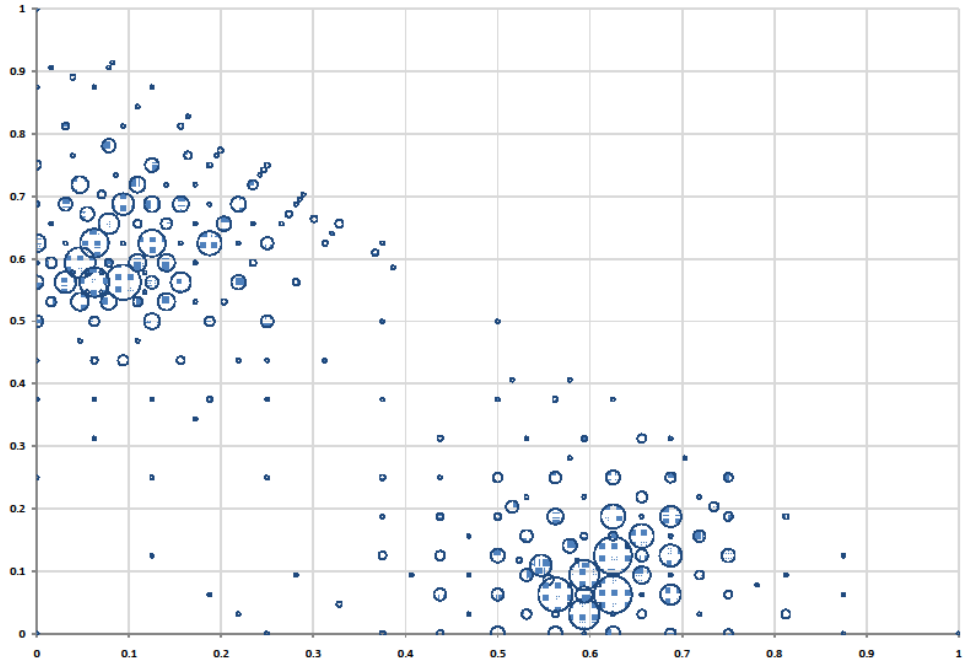


Fig. 56 : *Ensemble des points de simulation de* F *pour un niveau de bruit* σ_B *= 0.2, critère de scission no 1, pour une variabilité (*σ_K *,* α *) = (1, 1.5), et un nombre minimal de tirage* n_0 *= 10. La surface des bulles est proportionnelle au nombre de tirages de* F *en chaque point.* $\vec{\rho}$ *= (0.10, 0.011, 1.5E-05, 1.6E-02)*

II-4-3 Comportement de l'algorithme avec retirages possibles

Les deux graphiques 54 et 55 illustrent les retirages qui peuvent s'opérer sur des sommets déjà explorés, afin d'améliorer la connaissance de la fonction f en ces points. Pour ces figures, nous avons utilisé le critère de scission n° 2 envisageant l'amélioration du potentiel après ajout de n_0 simulations (cf. § II-3-1 de ce chapitre). Lorsque le niveau de bruit est σ_B = 0.1, le graphique 54 montre que les retirages ont essentiellement lieu autour des points solutions supposés, lorsqu'il n'est plus seulement nécessaire d'avoir une idée de la zone où se situe un minimiseur, mais qu'il faut également estimer avec précision la valeur m^* de l'optimum atteint. Lorsque le bruit augmente, cela tend naturellement à disperser un peu les points d'exploration, et favorise le retirage, dans la mesure où l'incertitude en certains points, du fait du bruit, devient plus forte que l'incertitude liée à la variation de la fonction entre les points de la grille.

Le graphique 56 illustre un choix de scission sur critère n° 1 comparant directement erreur de grille et erreur d'estimation (cf. § II-3-1 de ce chapitre) : elle donne une idée de l'impact du choix du critère de scission, qui conduit à un comportement proche pour les deux critères. Les retirages sur des sommets existants se font lorsque la présence du minimum est suffisamment vraisemblable et que l'incertitude pesant sur les sommets de la zone doit être diminuée.

Pour les deux critères abordés ici, il faut enfin noter que pour un niveau de bruit nul, aucun retirage n'intervient (les graphiques illustrant ce résultat, qui conduisent au graphique 49, ont toutefois été omis). Cela est logique dans la mesure où ces retirages n'apporteraient rien à la connaissance de la fonction f.

Pour l'algorithme avec retirages possibles, le comportement attendu est le suivant, conforme à ce qui a été observé :

– Lorsque le bruit est faible, ou à plus forte raison lorsque F est déterministe, la seule incertitude réside dans le comportement de la fonction entre les points, et l'algorithme crée systématiquement de nouveaux points de tirage.

– En présence d'un bruit, les points de retirage se concentrent dans les zones petites, généralement proches d'un point solution. Pour ces points, la réduction de l'erreur de grille, faible sur de courtes distances, est moins prioritaire que la réduction de l'erreur d'estimation.

	$\sigma_B = 0$ scission syst.		$\sigma_B = 0.1, \sigma_e = 0$ scission syst.		$\sigma_B = 0.1, \sigma_e$ estimé scission syst.		$\sigma_B = 0.1, \sigma_e$ estimé avec retirages.	
	σ_K	α	σ_K	α	σ_K	α	σ_K	α
n=10	0.635	1.505	0.475	1.212	0.474	1.214	0.449	1.277
n=100	0.946	1.592	0.392	1.063	0.645	1.374	0.947	1.585
n=1000	1.074	1.566	0.112	0.515	0.554	1.221	0.599	1.203

Tab. 34 : *Coefficients σ_K et α estimées par maximum de vraisemblance, en présence ou non d'un bruit σ_B, à partir de n points explorés*

II-4-4 Estimation des paramètres (σ_K, α)

Le tableau 34 représente les estimations des coefficients σ_K et α, opérées à partir d'un nombre n de points explorés, pour des données soumises à un bruit σ_B. L'estimation est réalisée par maximum de vraisemblance, comme indiqué au paragraphe II-2-3 de ce chapitre, à l'aide d'un point fixe qui a convergé dans toutes les situations testées. Ici, seuls les segments délimitant les zones ont été retenus pour l'estimation, et non pas les segments joignant deux points explorés de zones distinctes. Les zones en question ont été obtenues par l'algorithme avec des paramètres initiaux (σ_K, α)= (0.3, 0.9).

Dans la première colonne de ce tableau 34, lorsque σ_B = 0 (et en conséquence σ_e = 0), nous pouvons constater une relative stabilité des paramètres estimés, même dans le cas où l'estimation se base sur un nombre restreint de segments (lorsque n = 10 et en l'absence de considération de segments inter-zones). Dans la deuxième colonne, en italique, lorsque σ_B = 0.1 et pour tout point $\sigma_e(\theta)$ = 0, l'estimation est menée comme si l'estimateur $\hat{f}$ correspondait à la fonction f aux points explorés. Cette colonne 2 est donnée à titre indicatif, dans la mesure où il est parfaitement possible de tenir compte des erreurs d'estimation aux points explorés (cf. colonnes 3 et 4).

La non prise en compte des erreurs d'estimation explique le coefficient α inférieur dans le cas bruité, qui correspond bien à une incertitude accrue sur les petites distances, du fait du bruit supporté par $\hat{f}$. Ce phénomène est d'autant plus marqué que les zones sont petites (ici lorsque n est grand). Dans le cas bruité, le coefficient α inférieur est toutefois compensé par une pente σ_K inférieure. Enfin, dans les troisième et quatrième colonnes, l'estimation est menée sur des données bruitée, en tenant compte cette fois du bruit σ_e aux points explorés. La troisième colonne présente une estimation issue d'un découpage par scission systématique des zones, tandis que la quatrième présente une estimation issue de l'algorithme avec retirages possibles. Les paramètres observés sur ces colonnes 3 et 4 peuvent varier un peu, notamment lorsque le nombre de segments utilisés pour l'estimation est réduit.

Toutefois, dans les observations que nous avons pu mener, une valeur moins élevée de α (volatilité supérieure sur de courtes distances) est systématiquement compensée par une valeur moins élevée de σ_K (volatilité globale inférieure) : nous aborderons ce point par l'observation des écarts-types de la variable aléatoire IR_α dans le graphique 58.

Le comportement de l'algorithme initialisé avec les différents paramètres optimaux obtenus est resté assez stable dans les exemples que nous avons testés. C'est ce qu'indique le graphique 57, qui présente les points d'exploration obtenus, lorsque σ_B = 0.1, pour l'algorithme initialisé avec les paramètres estimés hors bruit pour n = 1000, (σ_K, α)= (1.0743, 1.5665) (à gauche) ou estimés en présence de bruit (σ_K, α)= (0.5545; 1.2209) (à droite). Un coefficient α supérieur permet de focaliser les recherches un peu plus rapidement sur de petites zones, mais la différence de valeur entre les coefficients α testés est ici trop ténue pour que l'effet global sur le comportement de l'algorithme en soit beaucoup affecté.

Enfin, nous avons déterminé, indépendamment de l'algorithme proposé, l'écart-type empirique de la variable aléatoire IR_α :

$$IR_\alpha(Z) = \frac{f(\theta_1) - f(\theta_2)}{d(\theta_1,\theta_2)^\alpha}$$

θ_1,θ_2 vecteurs aléatoires distincts issus de tirages indépendants, uniformes sur Z.

Le détail de la cette variable aléatoire et des tirages uniformes sur une zone Z est donné dans la sous-section II-2-3 de ce chapitre. L'écart-type de cette variable aléatoire fournit des informations sur la variation des pentes observées sur la zone en fonction de la distance considérée. Comme nous pouvons le voir dans le graphique 58, ces écarts-type sont légèrement supérieurs dans le cas de la demi-zone Z_1 correspondant à la partie de Θ située au dessus de la droite d'équation $y = x$.

Cela est logique dans la mesure où f a la forme d'une cuvette sur cette zone et non plus de deux cuvettes, les pentes sont ici en moyenne un peu plus abruptes que lorsque nous relions deux points d'une même cuvette, plutôt que deux points de cuvettes différentes. Nous pouvons remarquer que les ordres de grandeurs trouvés par maximum de vraisemblance sont tout à fait conformes à l'écart-type empirique de IR_α, notamment de celui de $IR_\alpha(Z_1)$, dans la mesure où dès la première itération de l'algorithme, le simplexe initial est scindé en une zone Z_1 et sa zone complémentaire.

La variable IR_α n'est pas utilisée par l'algorithme, mais fournit une indication sur l'erreur faite *a priori* sur la régularité de la fonction objectif selon le modèle proposé. Même il ne s'agit pas ici de fournir une étude exhaustive de la distribution de IR_α, qui dépend naturellement de la fonction objectif utilisée, il nous a paru intéressant de montrer quelle pouvait être la nature de cette erreur de modèle sur les données ici testées. Le résultat de cette analyse apparaît dans le graphique 59.

Sur notre fonction objectif, la distribution a posteriori de IR_α s'est révélée finalement assez proche de la distribution *a priori* supposée gaussienne, notamment pour $\alpha = 1.2$. Pour des α plus élevés, la présence de nombreuses valeurs très élevées se traduit par une légère sous-évaluation de la fréquence des pentes très élevées, et une sur-évaluation de la fréquence des pentes moyennes : la distribution observée a posteriori semble leptokurtique. Il est également possible que la distribution de IR_α évolue au fur et à mesure des découpages.

Ces résultats indiquent qu'il appartient à l'observateur d'intégrer une éventuelle prudence, en surestimant par exemple le coefficient σ_K ou en sous-estimant le coefficient α, et qu'il est aussi envisageable de modifier a posteriori la distribution supposée des pentes, ou le *kurtosis* de la distribution, l'hypothèse gaussienne pouvant servir de distribution a priori lors d'une inférence bayesienne.

Un écart important peut également indiquer une erreur dans le choix du coefficient α, le modèle devant conduire à un IR_α proche d'une loi normale pour un α à déterminer, non pour tous. Remarquons toutefois que, convolués avec l'erreur d'estimation en chaque point, les écarts observés peuvent en partie s'estomper. D'autre part, si la hiérarchie entre les zones les

plus susceptibles de contenir un optimum global peut se trouver affectée par le décalage observé, elle n'est pas non plus radicalement remise en cause, le coefficient d'aplatissement de la distribution utilisé ayant peu d'incidence pour des valeurs de f proches de la valeur estimée de l'optimum global (pour une zone dont les sommets conduiraient à $\hat{f}(\theta_i) = \hat{m}*$, le potentiel en tout point de la zone serait égal à 1, quel que soit le coefficient d'aplatissement de la distribution utilisée) : le potentiel est utilisé pour opérer un arbitrage entre différents points d'exploration possibles, non pour prévoir précisément le comportement de la fonction en ces points.

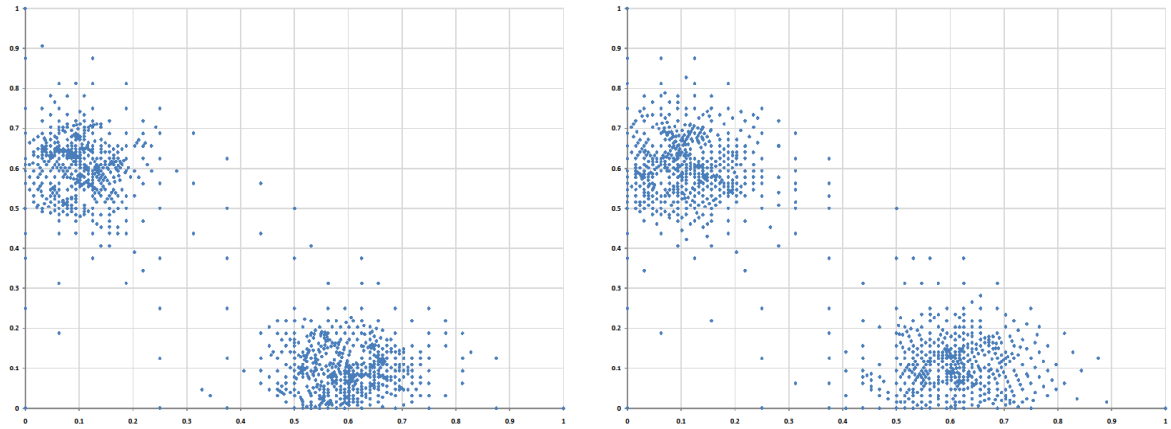


Fig. 57 : *Points d'explorations obtenus, lorsque σ_B = 0.1, avec les paramètres estimés hors bruit, (σ_K , α)= (1.0743, 1.5665) (à gauche) ou estimés en présence de bruit (σ_K , α) = (0.5545; 1.2209) (à droite)*

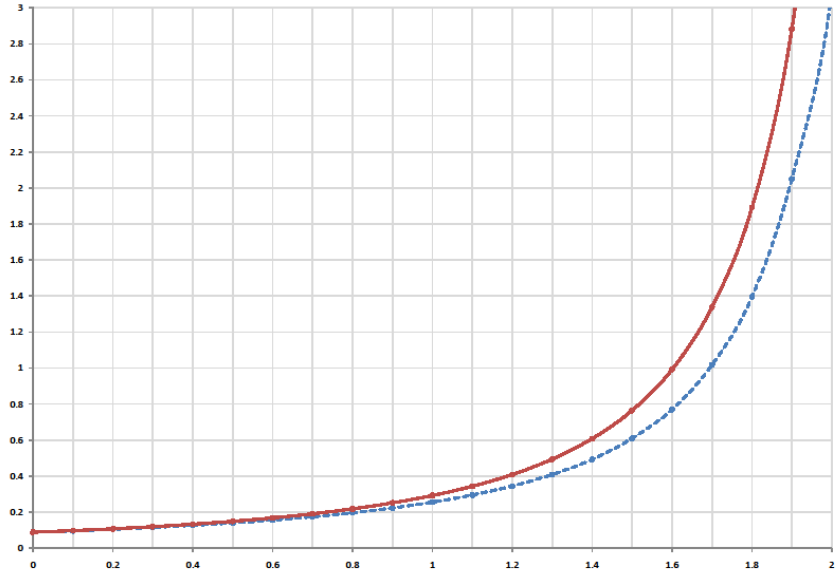


Fig. 58 : *Ecart-type empirique de $IR_\alpha(Z)$ obtenu pour la zone Z_0 correspondant au simplexe orthogonal standard initial (en pointillé) et pour la demi-zone Z_1 correspondant à la partie de Z_0 située au dessus de la première bissectrice (trait plein)*

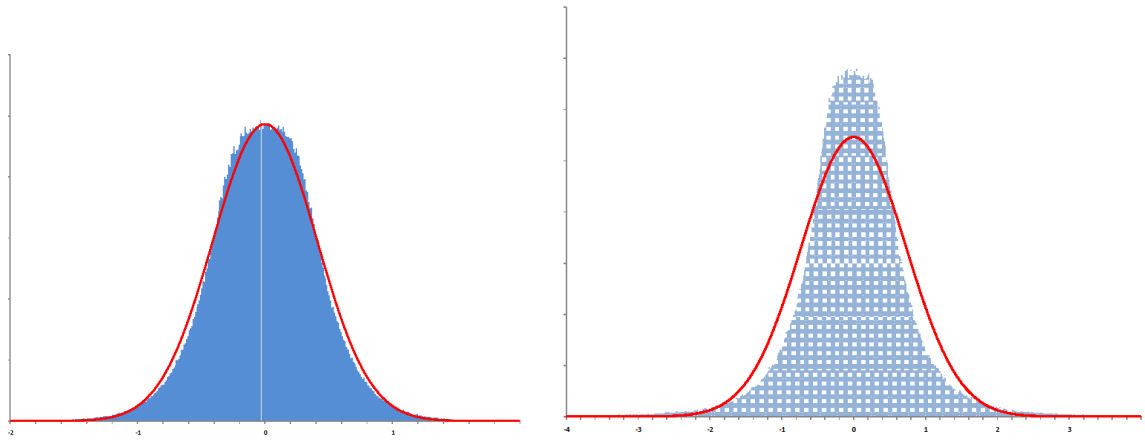


Fig. 59 : *Distribution a priori de* IR_α *(courbe gaussienne continue au premier plan) et histogramme a posteriori, pour la demi-zone* Z_1 *obtenue après la première itération de l'algorithme. Cas* α *= 1.2 (à gauche) et* α *= 1.5 (à droite). Les écarts-type empiriques obtenus à partir de l'histogramme sont respectivement* $\sigma = 0.41$ *et* $\sigma = 0.73$

II-4-5 Critères de convergence

critères de convergence calculés Le critère $\rho_1(s)$ est défini dans la sous-section II-2-6 de ce chapitre. Il s'appuie sur une zone de confiance proposée pour l'ensemble des solutions, zone comprenant un ensemble de points de potentiel supérieur à un seuil s. Le critère $\rho_1(s)$ indique la plus grande distance possible entre un point de cette zone et le point solution le plus proche : il peut donc s'interpréter comme un rayon maximal de la zone de confiance depuis chaque point solution. Le critère ρ_4, dépend quant-à-lui d'un seuil η qui sera, dans les applications numériques présentées, toujours fixé ainsi $\eta = 0.01$. Ce critère indique la plus grande probabilité (selon le modèle choisi), qu'un minimum plus petit que $(m^* - \eta)$ soit observé en un point θ (cf. sous-section II-2-6 pour une définition précise).

Dans les applications numériques, pour le calcul des critères ρ_1 et ρ_4, l'ensemble des points candidats Θ_0 a été construit sans opérer de nouveaux tirages de F, en piochant a posteriori 50 points dans chaque zone (de façon uniforme, cf. Smith et Tromble [2004]), et en calculant leur potentiel en fonction des points explorés aux sommets de chaque zone. Pour le critère ρ_1, l'ensemble des sommets proposés $\hat{S}_{m^*,s}$ a été extrait de Θ_0 en ne retenant que les points dont le potentiel était supérieur au seuil s fixé. Il est enfin rappelé que le critère ρ_2 est construit sur l'ensemble Θ_e des sommets explorés, et le critère ρ_3 sur l'ensemble des zones construites (cf. § II-2-6) :

$$\hat{S}_{m^*,s} = \{\theta \in \Theta_0, \beta_z(\theta) \geq s\}$$

$$\rho_1(s) = d^H\left(\hat{S}_{m^*,s}, S_{m^*}\right)$$

$$\rho_2 = \sup_{y \in S_{m^*}} \inf_{x \in \Theta_e} d(x,y)$$

$$\rho_3 = \max_{Z \in z_n} \beta(Z)$$

$$\rho_4 = \max_{\theta \in \Theta_0} \beta_z^{(m^*-\eta)}(\theta)$$

Fonction objectif non bruitée Bien que cette situation idéale ne soit pas celle nous désirons traiter dans la pratique, nous avons tout d'abord vérifié numériquement la convergence de l'algorithme sur la fonction test non bruitée, dans le cas $\sigma_B = 0$.

Les résultats numériques obtenus pour les critères de convergence apparaissent dans le tableau 35. Du fait de l'absence de bruit, le minimum estimé $\hat{m}*$ était supposé ne pas souffrir d'erreur d'estimation. Nous avons à titre indicatif ajouté dans le tableau 35 une ligne indiquant les résultats obtenus tenant compte d'une erreur de grille σ_{m^*} pour ce minimum, fonction du diamètre d_* de la zone sur laquelle il était présent, $\sigma_{m^*} = \sigma d_*^{\alpha}$. La localisation du minimum est alors plus dispersée, conduisant alors à une exploration accrue dans une zone plus étalée autour des minima trouvés, au détriment de la vitesse de convergence.

Nous pouvons notamment constater dans le tableau 35 le très bon comportement de l'algorithme en l'absence de bruit, avec des points explorés à une distance d'ordre 10^{-5} de chacune des solutions, et des zones de confiance proposées dans un rayon d'ordre $\rho_1(5\%) \cong 4.10^{-3}$ ou autour de $\rho_1(90\%) \cong 3.10^{-4}$ ces solutions. La zone de confiance proposée est naturellement plus réduite dans le cas où s est élevé.

	$\rho_1(s)$	ρ_2	ρ_3	ρ_4
s =90 %, σ_{m^*} =0	2.53 E-04	3.45 E-05	2.43 E-08	0
s =5 %, σ_{m^*} =0	4.32 E-03	3.45 E-05	2.43 E-08	0
s =5 %, σ_{m^*} >0 (erreur de grille)	6.46 E-03	1.38 E-04	2.28 E-07	0

Tab. 35 : *Critères de convergence obtenus par l'algorithme lorsque $\sigma_B = 0$, pour $\alpha = 1$ et $\alpha = 1.6$, avec ou sans prise en compte d'erreur de grille pour $m*$*

	$\rho_1(s)$	ρ_2	ρ_3	ρ_4
cas a, σ_B =0.1, sans retirage	7.83 E-02	2.21 E-03	1.32 E-06	3.24 E-02
cas b, σ_B =0.1, retirages	0.11	8.84 E-03	1.82 E-05	3.99 E-02
cas c, σ_B =0.1, retirages	5.92 E-02	2.21 E-03	6.14 E-06	4.58 E-02

Tab. 36 : *Critères de convergence obtenus pour les algorithmes avec ou sans retirage*

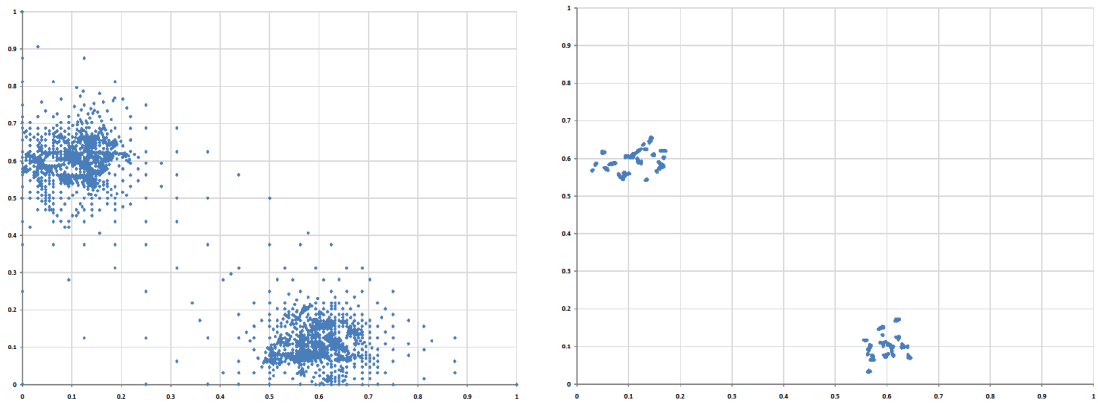


Fig. 60 : *Points d'explorations obtenus pour l'algorithme à scission systématique, cas a, lorsque* σ_B *= 0.1, avec les paramètres (*σ_K *,*α*) = (0.6, 1.2) (à gauche). Zone finalement retenue* $S_{m^*,5\%}$ *avec un seuil de 5 % (à droite)*

Fonction objectif bruitée Nous avons comparé les résultats obtenus par l'algorithme à scission systématique, ainsi que par l'algorithme à retirage pour les critères de scission n° 1 et n° 2. Les graphiques 60, 61 et 62 illustrent cette comparaison. Ces figures ont été obtenues pour un niveau de bruit σ_B = 0.1, des paramètres de régularités (σ_K , α) = (0.6, 1.2), pour n = 2000 points de tirage ou retirage, avec un seuil minimal de n_0 = 10 tirages par points. Dans ces figures apparaissent les nuages de points de tirage, ainsi que les zones de confiance proposées $S_{m^*,5\%}$. Du fait de l'absence de prise en compte d'erreur de grille pour m^*, et du fait du seuil s utilisé, il est possible qu'un point solution soit très proche mais en dehors d'une zone de confiance.
Les trois algorithmes se comportent correctement, et conduisent à la proposition de zones de confiance dans le voisinage direct des points solutions. Du fait du bruit frappant la fonction f, nous observons que des zones très proches peuvent être tantôt exclues, tantôt englobées dans $S_{m^*,s}$. Les zones de confiance ainsi bâties sont donc très fractionnées, mais bien localisées. Le critère n° 2 basé sur l'estimation des potentiels après ajout de n_0 tirages, a conduit à une zone de confiance légèrement moins dispersée autour des solutions connues.

Les résultats sont récapitulés pour ces mêmes tirages dans le tableau 36. Au regard du critère $\rho_1(5\%)$, l'algorithme à retirage fournit sur ces données les meilleurs résultats pour le critère n° 2.

Ces résultats peuvent néanmoins varier d'une exécution à l'autre du fait du caractère aléatoire du bruit pesant sur f et du caractère stochastique de l'algorithme. L'étude de la distribution de $\rho_1(s)$ constitue à cet égard un champ d'investigation intéressant.

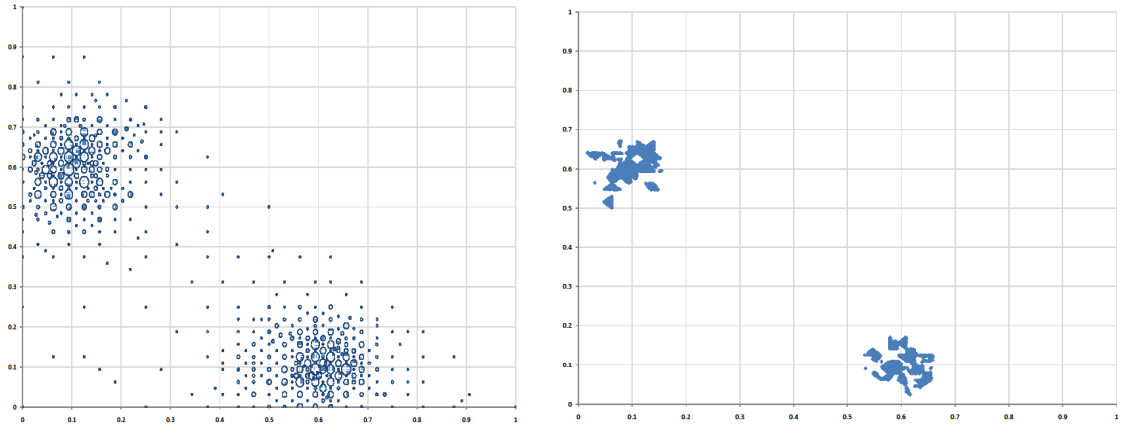


Fig. 61 : *Points d'explorations obtenus pour l'algorithme avec retirage, cas b, critère de scission n° 1 de comparaison des erreurs de grille et d'estimation, lorsque* σ_B *= 0.1, avec les paramètres* (σ_K, α)*= (0.6, 1.2) (à gauche). Zone finalement retenue* $S_{m^*,5\%}$ *avec un seuil de 5 % (à droite)*

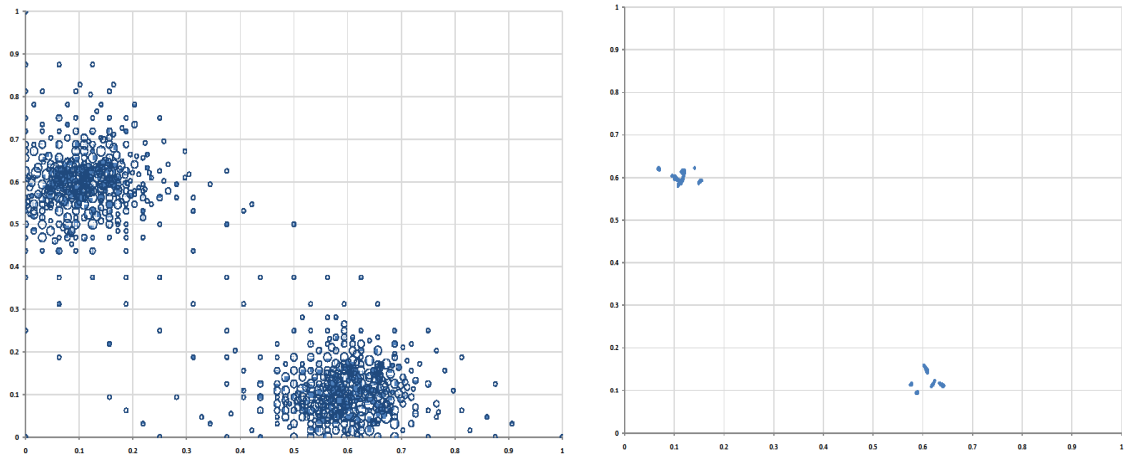


Fig. 62 : *Points d'explorations obtenus pour l'algorithme avec retirage, cas c, lorsque* σ_B *= 0.1,critère de scission n° 2 considérant l'ajout de* n_0 *observations, avec les paramètres* (σ_K, α)*= (0.6, 1.2) (à gauche). Zone finalement retenue* $S_{m^*,5\%}$ *avec un seuil de 5 % (à droite).*

Priorité	bruit σ_B	(σ_K, α)	$\rho_1(s)$	ρ_2	ρ_3	$\rho_4(\eta)$
absolue	0.1	(0.6, 1.2)	9.23 E-02	2.21 E-03	6.46 E-08	1.08 E-02
relative	0.1	(0.6, 1.2)	8.87 E-02	2.76 E-03	3.78 E-07	9.57 E-03
absolue	0	(0.6, 1.2)	3.61 E-02	5.52 E-04	1.17 E-06	0
relative	0	(0.6, 1.2)	6.18 E-02	5.52 E-04	6.77 E-06	8.05 E-05
absolue	0.1	(1, 1.6)	9.79 E-02	2.21 E-03	2.04 E-07	4.72 E-02
relative	0.1	(1, 1.6)	8.69 E-02	1.61 E-03	1.01 E-06	1.79 E-02
absolue	0	(1, 1.6)	1.26 E-03	3.45 E-05	3.17 E-09	0
relative	0	(1, 1.6)	2.57 E-03	3.45 E-05	2.57 E-08	0

Tab. 37 : *Comparaison des critères de convergence pour les choix priorité relative (exploration d'une zone avec une probabilité proportionnelle à son potentiel) ou priorité absolue (exploration de la zone de meilleur potentiel)*

Impact de la priorité Nous avons omis ici les graphiques qui permettaient de discuter du choix de γ , paramètre présenté dans la sous-section II-2-2. Nos observations nous conduisaient à un impact limité de ce paramètre. Dans les lignes suivantes, nous allons néanmoins évoquer plus précisément le choix de la priorité accordée à la zone de meilleur potentiel.

Jusqu'à présent, la zone à subdiviser ou explorer était choisie avec une probabilité proportionnelle au potentiel de la zone (choix qualifié ici "*priorité relative*"). Il est également envisageable de ne plus choisir la prochaine zone de façon stochastique, mais de choisir de façon déterministe zone de meilleur potentiel (choix qualifié ici "*priorité absolue*"). Cela revient à traiter une question déjà abordée : choisir entre explorer davantage la fonction, ou privilégier la vitesse en explorant en priorité une zone prometteuse, au risque d'ignorer un optimum global.

Les paramètres (σ_K, α) permettent justement d'opérer cet arbitrage. Si ces paramètres sont en partie estimés, ce dernier choix de priorité *relative* versus *absolue* conduit à un arbitrage naturellement différent. Dans les applications numériques opérées, ce choix n'a pas eu beaucoup d'incidence, notamment en présence de bruit.

Le tableau 37 récapitule les résultats obtenus avec le choix de priorité relative ou le choix absolue. Nous nous sommes placé pour cette application dans un cadre de scission systématique (pas de retirages), avec un nombre de tirages n = 2000. Pour le critère ρ_1, nous avons bâti l'ensemble solution proposé à l'aide d'un seuil s égal à 10 % du meilleur potentiel observé, en conservant ainsi une part notable des solutions possibles.

Rappelons que le choix d'un seuil supérieur conduit à rétrécir les zones de confiance autour des optima trouvés, et peut améliorer de façon importante le critère de convergence ρ_1, comme nous pouvons le constater en comparant ces résultats avec ceux du tableau 36. Ce critère ρ_1 n'est donc pas comparable à d'autres critères qui seraient obtenus avec d'autres seuils. Par ailleurs l'augmentation du seuil s augmenterait le risque d'ignorer un optimum global, et le critère ρ_1 pourrait alors être brutalement dégradé.

Remarquons tout d'abord dans le tableau 37 la bonne exploration des voisinages de chacun des points solutions : dans tous les cas, après 2000 itérations, des points ont été tirés à une

distance ρ_2 inférieure à 3.10^{-03} de chacune des solutions, avec des zones de confiance d'un rayon ρ_1 raisonnable autour des solutions (qu'il est possible de réduire en augmentant le seuil s). En l'absence de bruit, des tirages sont obtenus, avec (σ_K , α)= (1, 1.6), à une distance très réduite d'environ 3.10^{-05} de chaque point solution.

Au vu des mesures effectuées dans ce tableau 37, le *choix priorité relative* ou *priorité absolue* a surtout un impact dans les situations non bruitées où les paramètres (σ_K , α)= (1, 1.6) privilégient une moindre exploration. Dans ce cas, la convergence est assez rapide, et le choix *priorité absolue* conduit à un zone de confiance de diamètre faible ρ_1= (1.2)**.**10^{-03}, plus faible que dans le cas *priorité relative*. Dans les autres cas, il apparaît que les mesures ne sont pas radicalement perturbées par ce choix, notamment dans le cas d'un bruit σ_B non nul.

II-4-6 Comparaison avec l'algorithme de Kiefer-Wolfowitz-Blum

Notre algorithme vise à opérer une optimisation à partir d'une fonction F bruitée, et il n'est pas en pratique possible d'éliminer ce bruit. Pour cette raison, nous ne comparerons pas notre algorithme avec les algorithmes d'optimisation classiques en l'absence de bruit. Nous avons donc choisi de comparer les résultats obtenus avec ceux que donnerait un autre algorithme d'optimisation de fonction bruitée.

L'algorithme utilisé pour la comparaison est un des algorithmes d'optimisation stochastique parmi les plus classiques : il s'agit de l'extension multidimensionnelle, proposée par Blum [1954] de l'algorithme de Kiefer et Wolfowitz [1952]. Rappelons tout d'abord que l'algorithme de Kiefer- Wolfowitz vise l'obtention d'un unique optimum, non la garantie d'absence d'autres optima. Tel que décrite dans Broadie et al. [2009], la version de l'algorithme de Blum consiste à déterminer une suite de points $\theta^{(1)}, \theta^{(2)}, \ldots$ convergeant vers la solution θ^*, supposée unique, minimisant $f(\theta) = E[F(\theta)]$. En dimension d = 2, si l'on note $\theta^{(k)} = \left(\theta_x^{(k)}, \theta_y^{(k)}\right), k \in \mathrm{N}$ la suite de points est telle que :

$$\theta_x^{(n+1)} = \theta_x^{(n)} - a_n \frac{F\left(\theta^{(n)} + c_n e_x\right) - F\left(\theta^{(n)}\right)}{c_n}$$

$$\theta_y^{(n+1)} = \theta_y^{(n)} - a_n \frac{F\left(\theta^{(n)} + c_n e_y\right) - F\left(\theta^{(n)}\right)}{c_n}$$

où e_x = (1, 0) et e_y = (0, 1), et où $\{a_n\}_{n\in \mathrm{N}}$ et $\{c_n\}_{n\in \mathrm{N}}$ sont des suites réelles décroissantes en n. Les contraintes auxquelles sont soumises ces constantes, ainsi que les conditions générales d'application et de convergence de l'algorithme, sont détaillées dans Broadie et al. [2009].

Remarquons qu'à chaque étape de l'algorithme, le budget de tirage de F est amputé de d + 1 = 3 tirages : d tirages pour estimer le gradient, ici en $\theta^{(n)} + c_n e_x$ et $\theta^{(n)} + c_n e_y$, et un tirage pour le nouveau point $\theta^{(n+1)}$. Pour un budget de n points de tirage de la fonction F supposée coûteuse, l'algorithme ne pourra pas faire appel à un nombre d'étapes supérieur à (n/3). A cet égard, cette version de l'algorithme, faisant appel à des différences finies sur un seul côté, dépense moins de tirages que d'autres versions pour l'estimation du seul gradient.

Par ailleurs, les suites $\{a_n\}$ et $\{c_n\}$ semblent délicates à choisir, et les écrits de Broadie et al. [2009] mentionnent une grande sensibilité du comportement de l'algorithme en fonction de ces choix. En l'absence supposée d'autres informations sur f, nous avons opté pour un choix par défaut classique de $a_n = 1/n$ et $c_n = 1/n^{1/4}$, choix proposé en première page dans l'article originel Kiefer, Wolfowitz [1952]. Les paramètres $a_n = a_0 / n$, $c_n = c_0 / n^{1/3}$, $n \geq 1$, présentés dans Broadie et al. [2009] conduisaient ici à des résultats décevants pour $a_0 = c_0 = 1$ et les quelques choix alternatifs testés, mais nous n'avons pas cherché spécifiquement à trouver a posteriori les meilleurs paramètres, pour une fonction f supposée inconnue.

En conséquence, les résultats numériques obtenus ici sont représentatifs d'un type de trajectoire classiquement obtenue lors d'une descente stochastique de gradient, mais certainement améliorables en terme de vitesse de convergence, les suites $\{a_n\}$ et $\{c_n\}$ pouvant être modifiées à cette fin.

Dans la plupart des illustrations précédentes, pour l'algorithme à scission systématique, n_0 = 10 tirages étaient opérés en chacun des points explorés, pour un budget de 2000 points d'exploration. Le budget global de tirage de F était donc de 2000 n_0 tirages. Le graphique 63 rend compte des points de tirage obtenus par l'algorithme de Kiefer-Wolfowitz- Blum, pour 2000 points de tirage d'une fonction F soumise à un bruit $\sigma_B / \sqrt{n_0}$ (figure de gauche), ou pour 2000 n_0 tirages d'une fonction F soumise à un bruit σ_B (figure de droite). Le nombre global de tirage correspond ainsi à celui des illustrations précédentes. Le point initial a été tiré aléatoirement, de façon uniforme, sur le simplexe initial (cf. Smith et Tromble [2004]).

Sur le graphique 63, nous constatons tout d'abord la présence de trois séries de points, correspondant à l'ensemble des points suggérés, ainsi qu'aux points décalés verticalement et horizontalement pour l'estimation du gradient (points décalés respectivement vers le haut et vers la droite). Les points sont initialement assez espacés (a_n grand) puis sont de plus en plus proches au fil des tirages (a_n petits, lorsque nous sommes en principe proche d'une solution). Sur les données moins bruitées (figure de gauche), les gradients estimés sont moins erratiques, les directions choisies d'un point à l'autre étant plus stables.

Ce graphique 63 est essentiellement intéressant dans la mesure où il marque bien les différences d'approches et de finalité entre les algorithmes de descente stochastique de gradient et l'algorithme proposé dans le présent travail. Concernant l'algorithme de Kiefer-Wolfowitz- Blum, nous pouvons notamment faire les constats suivants :

– En dehors de l'espace occupé par les trois trajectoires qui se dessinent, la fonction F est très peu explorée sur le reste de la zone de recherche, et est donc susceptible d'ignorer des optima globaux. Cela est logique dans la mesure où nous n'exigeons normalement pas d'un algorithme d'optimisation locale qu'il fournisse un optimum global.

– Par ailleurs, même en présence supposée d'un unique optimum, la construction de zones de confiance pour l'optimum, à partir des trois trajectoires obtenues, semble ici délicate.

– Des points à l'extérieur du simplexe initial Z_0 ont pu être tirés, dans la mesure où $f(\theta)$ pouvait prendre ici des valeurs en dehors du simplexe initial, mais l'exploitation de cet algorithme supposerait l'introduction préalable de contraintes.

– L'algorithme se comporte ici de façon cohérente dans la mesure où la fonction f est suffisamment régulière sur une grande partie du simplexe initial (en dehors de la première diagonale). La situation serait bien différente en cas de non-dérivabilité de f sur l'essentiel du domaine Z_0.

S'agissant des critères de convergence, l'algorithme de Kiefer-Wolfowitz-Blum ne permet malheureusement pas *a priori* de partitionner la zone de recherche initiale Z_0 en plusieurs zones.

En conséquence, nous ne calculerons pas ici de potentiel. En supposant que Θ_e désigne l'ensemble des points de tirage obtenus, nous réutiliserons donc le critère ρ_2 :

$$\rho_2 = \sup_{y \in S_{m^*}} \inf_{x \in \Theta_e} d(x,y)$$

Ce critère permet de savoir s'il existe des points de tirage proches de chacune des solutions.

L'algorithme de Kiefer-Wolfowitz-Blum recherchant une unique solution est très fortement pénalisé par ce critère, car il peut converger vers une solution tout en ayant des points d'exploration très éloignés des autres solutions. Un second critère retenu spécifiquement pour cet algorithme sera :

$$\tilde{\rho}_2 = \inf_{y \in S_{m^*}} \inf_{x \in \Theta_e} d(x,y)$$

Ce dernier critère indique juste s'il existe des points de tirage proches de l'une des solutions. Les résultats obtenus sont indiqués dans le tableau 38, qui correspond aux courbes du graphique 63.

D'une part, l'algorithme de Kiefer-Wolfowitz-Blum ne recherche qu'un optimum local. Nous devrions donc obtenir un critère ρ_2 de convergence vers un optimum global très dégradé par rapport à l'algorithme de scission systématique, ce qui est bien le cas.

D'autre part, la recherche d'un optimum local doit intuitivement être moins coûteuse que la recherche d'un optimum global. Les points explorés près d'une unique solution devraient être plus proches de la solution pour un algorithme d'optimisation locale et nous devrions obtenir un critère ρ_2 meilleur. Ce n'est pas le cas ici : les mesures de ρ_2 obtenues avec l'algorithme à scission systématique ont conduit (cf. tableau 37) à des points toujours tirés à une distance ρ_2 inférieure à 3.10^{-3} de chaque solution. Or, la distance $\tilde{\rho}_2$ à la solution la mieux explorée est encore inférieure à ρ_2. L'algorithme de Kiefer-Wolfowitz Blum a conduit quant-à-lui à des résultats de l'ordre de 0.2 pour la distance ρ_2, et 3.10^{-2} pour la distance $\tilde{\rho}_2$. Notre algorithme a donc été ici localement et globalement plus performant. Ce dernier constat est sans doute lié à l'absence d'optimisation spécifique des séquences $\{a_n\}$ et $\{c_n\}$, qui reste toutefois délicate à opérer *a priori* sur une fonction supposée inconnue.

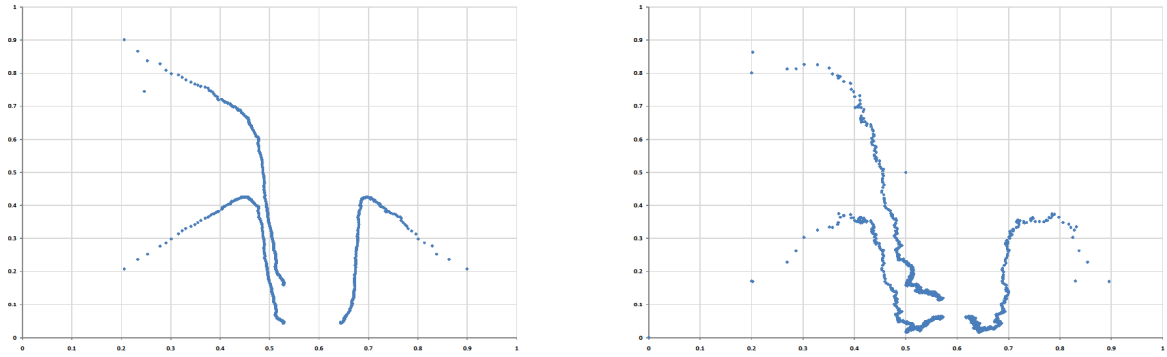


Fig. 63 : *Points d'explorations obtenus pour l'algorithme Kiefer-Wolfowitz-Blum, pour 2000 tirages avec* $\sigma_B = 0.1/\sqrt{10}$ *(à gauche), ou 20000 tirages avec* $\sigma_B = 0.1$ *(à droite)*

	ρ_2	$\tilde{\rho}_2$
Cas n=2000, σ_B=0.1/√10	0.2058	6.35 E-02
Cas n=20000, σ_B=0.1	0.2248	3.34 E-02

Tab. 38 : *Critères de convergence obtenus avec l'algorithme de Kiefer-Wolfowitz (pour des séquences* $\{a_n\}$ *et* $\{c_n\}$ *non spécifiquement optimisées)*

II-5 Conclusion

Nous avons présenté un algorithme permettant de rechercher des paramètres de $\Re^d$ conduisant à minimiser, globalement, une fonction réelle bruitée. L'algorithme permet également de construire des zones de confiance autour des paramètres solutions supposés. L'approche retenue s'appuie sur la définition du potentiel d'une fonction en un point, et sur la mesure de l'incertitude frappant la fonction objectif entre les points déjà explorés. Cette incertitude découle d'une part des *erreurs d'estimation* aux point explorés, d'autre part des *erreurs de grille* liées à la distance entre un point considéré et les points explorés alentours.

L'algorithme permet de moduler facilement, au moyen des deux paramètres σ_K et α, la réitération de tirages dans un voisinage des points solutions découverts, ou à l'inverse l'exploration de la fonction dans des zones encore peu explorées. En outre, une procédure d'estimation de la régularité fonction objectif f permet d'ajuster les valeurs de ces deux paramètres.

Pour des dimensions pas trop élevées, une étude a également été effectuée (cf. Rullière et al. [2010]). Les résultats obtenus montrent que le comportement de l'algorithme est assez convaincant sur de nombreux points, comme sa capacité d'exploration globale du domaine, sa capacité à privilégier les zones proches des optima, sa faible exigence sur la régularité de la fonction objectif, son adaptation à un bruit nul comme à un bruit élevé. Toutefois, lorsque la dimension du problème devient grande, l'algorithme conduit à des résultats proches de ceux obtenus par une exploration uniforme du domaine initial. L'analyse détaillée de la vitesse de convergence de l'algorithme et l'amélioration éventuelle de cette vitesse en grande dimension constituent une suite logique de cette étude.

Conclusion du chapitre 2

Le paysage des régimes de retraite en France est diversifié. Une panoplie large de contraintes et d'objectifs peut être rencontrée en pratique. Dans ce cadre, nous avons proposé une modélisation du régime-type de retraite tout en étudiant certains critères d'allocation stratégique d'actifs (en fonction du type du régime : provisionné, partiellement provisionné, etc.). La stratégie *Fixed-Mix*, une des référence en matière d'allocation d'actifs, a été ensuite explorée et mise en œuvre dans le cas d'un régime de retraite partiellement provisionné. Le souci de réduction des temps de calcul lors de cette mise en œuvre nous a conduits à proposer un algorithme d'optimisation numérique « par exploration sélective ».

Il s'agit d'un algorithme qui vise à montrer l'intérêt et la faisabilité technique d'une grille à pas variable, susceptible de répondre au problème du choix entre exploration et connaissance d'une fonction aléatoire. L'approche proposée ici se voulait simple, avec notamment la définition d'un potentiel directionnel pour une unique dimension, puis l'agrégation de potentiels directionnels sur l'ensemble des dimensions. Il est possible que de nombreux autres choix puissent être faits pour la mesure du potentiel d'un point et d'une zone. Les techniques d'agrégation et de tirages peuvent également être amendées de façon à améliorer l'efficacité de l'algorithme sur des architectures parallèles.

Dans le chapitre suivant, nous essayons d'étudier, sur un autre plan, le problème d'allocation stratégique d'actifs. L'objectif sera, cette fois-ci, de modifier la stratégie d'allocation et d'étudier l'effet sur le résultat final. En particulier, nous explorons différentes techniques permettant de nous positionner dans un contexte d'allocation dynamique, différent de celui considéré jusque là avec l'allocation à poids constants (*Fixed-Mix*).

Chapitre 3 : Allocation stratégique d'actifs et modèles d'ALM dynamique

Dans le chapitre précédent, le cadre général pour la mise en place d'un modèle de gestion actif-passif basé sur la stratégie *Fixed-Mix* a été présenté : ce type de modèle suppose un rebalancement périodique des actifs, de façon à garder constante la composition du portefeuille financier au début de chaque sous période. Même si la stratégie *Fixed-Mix* a le mérite de prendre en compte l'aspect aléatoire des scénarios économiques et financiers, cette stratégie peut être critiquée du fait qu'elle ne suppose pas une flexibilité des positions futures sur le marché face aux différents scénarios projetés de rendement. Plus précisément, cette approche ne permet pas d'utiliser complètement l'information qui arrive au fur et à mesure dans le temps pour optimiser le choix de l'allocation stratégique (cf. Zenios [2007]).

Une autre approche, tournée vers le futur (« *forward-looking* ») et permettant d'intégrer l'aspect dynamique dans la composition du portefeuille serait « *a priori* » d'une importance majeure. L'élément clé de distinction entre les deux stratégies (dynamique et *Fixed-Mix*) se situe au niveau des compositions des portefeuilles testés par le moteur d'optimisation, tout au long de la période de projection.

Dans le cas de la stratégie *Fixed-Mix*, nous partons à t=0 d'un vecteur de poids unique (ex. 40 % actions et 60 % obligations) et constant (puisque nous supposons un rebalancement périodique, par exemple annuel, de la composition du portefeuille vers ce vecteur). Nous calculons *in fine* (ex. à t=10 ans) les indicateurs risque/rendement correspondant. Ce calcul est appliqué pour différents vecteurs. Le but est de comparer les indicateurs relatifs aux différents vecteurs de poids pour finalement choisir celui qui optimise les critères d'allocation retenus. L'allocation stratégique optimale correspond dans ce cas au vecteur de poids optimal.

Dans le cas d'une stratégie dynamique, nous nous autorisons la possibilité d'avoir des compositions intermédiaires de poids différentes (ex. entre t=0 et t=10 ans). L'ensemble de ces compositions (ou vecteurs) intermédiaires donnera lieu à ce que nous pouvons appeler un « chemin » d'allocation. L'objectif dans le cadre de la stratégie dynamique est de comparer les indicateurs risque/rendement obtenus par chaque « chemin » pour finalement retenir celui qui optimise les critères retenus d'allocation. Nous aurions dans ce cas un chemin optimal ainsi qu'une allocation stratégique optimale (correspondant au vecteur de poids de la date t=0).

L'allocation statique de type *Fixed-Mix* peut être considérée comme un cas particulier de l'allocation dynamique : pour cela, il suffit de la voir comme une allocation dynamique avec des vecteurs intermédiaires identiques au vecteur de poids initial. Nous remarquons également que pour chacune des deux stratégies évoquées ci-dessus, la composition intermédiaire du portefeuille est supposée être ajustée mais selon deux règles différentes : tandis qu'avec la stratégie *Fixed-Mix* cet ajustement a pour but de converger vers un vecteur constant (rebalancement), l'ajustement de poids dans la stratégie dynamique cible la convergence vers des vecteurs, *a priori*, différents du vecteur de poids initial (à t=0).

En pratique, l'ajustement intermédiaire des poids peut ne pas avoir lieu : par exemple nous pouvons considérer le cas où à t=1, nous recalculons des allocations optimales (compte tenu des nouvelles informations arrivées entre t=0 et t=1), et que ce nouveau calcul donne des résultats d'allocations optimales différentes. Dans ce cas, l'allocation optimale initiale (celle de t=0) est délaissée en faveur de celle calculée à t=1.

Dans ce chapitre, les modèles classiques d'ALM dynamiques sont explorés, notamment les techniques d'assurance de portefeuille (cf. Perold et Sharpe [1988]) et les techniques de programmation dynamique (cf. Cox et Huang [1989], Merton [1971]). A ce niveau, les techniques d'assurance de portefeuille basées sur la notion de CPPI ou *Constant Proportion Portfolio Insurance* (cf. Perold et Sharpe [1988]) sont mises en place et certains résultats relatifs à ce modèle sont étudiés. De même, les principes des techniques de programmation dynamique sont également mis en évidence.

Les principes des techniques de programmation stochastique sont ensuite présentés. Ces techniques sont adaptées dans un contexte d'allocation stratégique d'actifs et d'ALM d'un régime de retraite typique (partiellement provisionné). Nous proposons, au cours de la même étude, une nouvelle méthodologie de génération de scénarios économiques que nous appelons méthodologie « des quantiles de référence ». A travers une application numérique, nous comparons certains résultats relatifs aux deux approches d'allocation stratégique d'actifs : celle basée sur la stratégie *Fixed-Mix* et celle basée sur les techniques de programmation stochastique. Nous testons également la sensibilité de cette dernière approche par rapport au changement de certains de ses paramètres, toutes choses étant égales par ailleurs.

I- Allocation stratégique d'actifs dans le cadre des modèles classiques d'ALM dynamique

I-1 Présentation générale

La gestion dynamique de portefeuille sur le long terme reste un domaine de recherche relativement peu exploré, par comparaison avec l'importance des travaux déjà réalisés sur les aspects à court terme. Le dilemme qui s'impose à ce niveau est celui de la définition des limites de l'allocation stratégique, là où commence l'allocation tactique. Autrement dit, il paraît assez judicieux que le lecteur se pose la question suivante : par cet aspect dynamique ne dépassons nous pas le cadre d'une allocation stratégique pour être amenés à définir une allocation tactique ? La réponse est toutefois négative.

En fait, en gardant un horizon d'analyse de long terme avec toutes les contraintes liées aux flux de passif (voire celles sur les actifs), nous gardons à l'esprit la finalité d'une allocation de long terme, qui n'est autre que la définition d'une allocation stratégique. La différence cruciale sera qu'on se permet un ajustement intermédiaire des poids du portefeuille, dans le but d'optimiser notre fonction objectif finale et de profiter des éventuelles opportunités qui se manifestent sur le marché, au fur et à mesure que l'incertitude disparaît.

L'allocation tactique commence là où l'allocation stratégique s'arrête. L'allocation stratégique est, par définition, une allocation qui se base sur des anticipations sur le long terme des rendements et du risque. Ces anticipations sont formées à partir d'une variété de facteurs, tel que les rendements historiques et les anticipations des indicateurs macro-économiques sur le long terme (voire même les risques politiques associés à certains pays). L'allocation stratégique est déterminée de façon non fréquente (au maximum une fois par an). En pratique, des marges de poids par rapport à cette allocation stratégique sont souvent définies afin de permettre la mise en place des stratégies tactiques.

C'est à l'allocation tactique de repositionner le profil de rendement/risque de l'allocation stratégique afin de répondre à des variations intermédiaires, de court terme, sur les marchés (par exemple avec un horizon quotidien ou mensuel de prévision). Par exemple, nous

pourrons envisager le cas simplifié où les expositions sur les marchés sont revues, de façon tactique, selon la règle de gestion suivante : la réduction des poids des classes d'actifs où le risque a de forte chance de croître et le renforcement des classes d'actifs ayant un potentiel de rendement positif élevé. L'allocation tactique se base essentiellement sur des analyses court-termistes du profil rendement/risque de chacune des classes d'actifs.

Dans ce qui suit, nous présentons brièvement les modèles classiques d'ALM dynamiques, notamment les techniques d'assurance de portefeuille (cf. Perold et Sharpe [1988]) et les techniques de programmation dynamique (cf. Cox et Huang [1989], Merton [1971]). A ce niveau, les techniques d'assurance de portefeuille basées sur la notion de CPPI ou *Constant Proportion Portfolio Insurance* (cf. Perold et Sharpe [1988]) sont mises en place à titre illustratif.

I-2 Techniques d'assurance de portefeuille

I-2-1 Principe

L'objectif de l'assurance de portefeuille est de permettre aux investisseurs de participer à la performance du marché, tout en leur offrant une protection à maturité d'une proportion du capital initial investi. Cette technique peut s'inscrire parmi les stratégies de gestion dynamique d'un surplus dans le cadre de l'ALM.

Une des méthodes standards d'assurance de portefeuille est la méthode dite « de l'assurance de portefeuille à proportion constante » (en anglais « *Constant Proportion Portfolio Insurance* » ou CPPI, cf. Perold et Sharpe [1988], Hilli et al. [2007]). Cette stratégie est fondée sur une allocation dynamique entre un actif risqué (par exemple les actions) et un actif sans risque (par exemple le monétaire) afin de garantir à chaque date un niveau minimal prédéterminé pour le portefeuille.

Dans la terminologie de l'assurance de portefeuille, nous définissons couramment :

- « Le Plancher » : la valeur minimale du portefeuille acceptable pour l'investisseur à maturité,
- « Le Coussin » (analogue au surplus en ALM) : l'écart entre la valeur du portefeuille et le plancher garanti,
- « L'exposition » : la proportion du portefeuille qui est investie en actif risqué.

Dans l'approche CPPI, la proportion du portefeuille investie dans l'actif risqué varie proportionnellement au montant placé en actif sans risque, de manière à assurer à chaque date le niveau plancher. Mathématiquement, cela consiste à introduire le paramètre m (appelé « multiplicateur »), tel qu'à chaque date de rebalancement du portefeuille nous ayons la relation :

$$Exposition = m \times Coussin$$

En pratique le multiplicateur doit être ajusté en fonction des fluctuations du marché, tout en restant à l'intérieur d'une zone cible.

I-2-2 Illustration

I-2-2-1 *Contexte*

Nous supposerons que les salariés de certaines catégories professionnelles auront la possibilité d'épargner tout ou partie de leur compte épargne-temps dans un fonds garanti, leur assurant à l'échéance la récupération du capital investi ainsi qu'une participation à la performance des marchés au cours de la vie du fonds.

La vocation de ce portefeuille étant d'offrir aux souscripteurs un complément de retraite, l'échéance est donc ajustée en fonction de l'âge du client au démarrage du produit. La conception du mécanisme de ce type de fonds garanti est l'objet de la présente étude.

Les données de départ consistent, à ce stade, en des échéanciers prévisionnels (établis par des techniques actuarielles) des flux annuels versés par les souscripteurs. Ces échéanciers de flux sont répartis en fonction de la tranche d'âge à laquelle appartient le souscripteur fin 2010 par exemple.

La stratégie proposée consiste à gérer le fonds selon une extension de la méthode CPPI, en prenant en compte le passif constitué des flux annuels de versements de RTT (Réduction du Temps de Travail) de la part des investisseurs.

Nous considérerons par ailleurs cinq fonds (gérés indépendamment les uns des autres) respectivement associés aux tranches d'âge de l'investisseur en 2010, soit 25 ans, 35 ans, 45 ans et 55 ans. A chacun de ces fonds correspond ainsi un échéancier prévisionnel différent de flux de RTT jusqu'à la retraite.

Pour chacun des différents fonds, les hypothèses à la date de lancement sont les suivantes :

- La valeur de l'actif du portefeuille est initialisée avec un montant égal au premier versement de RTT des souscripteurs pour la tranche d'âge associée
- Le passif (montant garanti à l'échéance de la retraite) est initialisé comme le montant du premier flux versé, actualisé entre la date de lancement et l'échéance

Nous notons que le calibrage du taux utilisé pour l'actualisation du passif peut se faire en fonction des résultats de simulations *Monte Carlo* du portefeuille (comme vu dans la suite).

Le mécanisme des flux de trésorerie sera supposé être le suivant. A chaque versement annuel de flux de RTT :

- l'actif du fonds est augmenté du montant du flux versé
- Le passif est augmenté de la valeur actualisée du flux entre la date de versement et l'échéance.

De plus, à chaque date mensuelle :

- L'allocation de l'actif du fonds entre actif risqué (actions) et actif sans risque (monétaire) est ajustée en appliquant la stratégie CPPI
- Le passif du fonds du mois précédent est capitalisé (via le même taux que celui utilisé pour l'actualisation)

I-2-2-2 *Modélisation de l'actif*

L'actif de chaque fonds est supposé être constitué :

- d'un actif risqué (indice action), le montant investi constituant l'exposition de l'actif du portefeuille.
- d'une partie investie en monétaire, dont le rendement est égal au taux sans risque.

Indépendamment de la gestion CPPI, l'exposition du portefeuille est supposée évoluer suivant un processus stochastique de type Black&Scholes [1973] (avec un pas de discrétisation mensuel).

Ainsi, à chaque date *t* (fréquence mensuelle), l'exposition *E*(*t*) évolue selon le processus suivant :

$$E(t) = E(t-1) \times e^{\left[\left(\mu_m - \sigma_m^2 / 2\right) + \sigma_m \times W_t\right]}$$

Les paramètres μ_m et σ_m sont respectivement, la performance mensuelle et la volatilité mensuelle de l'indice action. Les valeurs mensuelles de la performance de l'indice et de sa volatilité se déduisent des valeurs annualisées μ et σ par les formules de conversion usuelles.

I-2-2-3 *Modélisation du passif*

Le passif initial des différents fonds est pris égal à la valeur du premier flux versé, actualisée entre la date de versement et la date d'échéance (départ à la retraite)

Le taux mensuel t_m d'actualisation du passif, qui conditionne le niveau de la garantie (donc également le processus de réallocation dynamique), peut être calibré de manière spécifique ou bien être déduit des autres paramètres du modèle.

$$P(t_0) = F(t_0) / (1 + t_m)^{n_m}$$

Où :

$P(t_0)$: le passif initial des différents fonds (ou le passif de la date de lancement)

$F(t_0)$: premier flux versé

n_m : nombre de mois jusqu'à l'échéance (entre la date de versement et la date de départ à la retraite)

A fréquence mensuelle (entre chaque date annuelle de versement de flux), le passif précédent est capitalisé suivant l'expression :

$$P(t) = P(t-1) \times (1 + t_m)$$

A chaque date annuelle de versement de flux, le nouveau flux versé (actualisé entre la date de versement et l'échéance) s'ajoute de plus au passif précédent :

$$P(t) = P(t-1) \times (1 + t_m) + F(t) / (1 + t_m)^{n_m}$$

Le nouveau passif (après versement du flux) est égal à la somme entre d'une part le passif du mois précédent multiplié par (1+ t_m) et d'autre part le flux versé actualisé entre la date de versement et la date d'échéance.

I-2-2-4 *Calibrage des paramètres de marché*

Les différents paramètres de marché du modèle à calibrer sont les suivants :

- Taux sans risque annualisé (évolution de la partie monétaire de l'actif) : r
- Performance annualisée de l'actif risqué (indice actions) : μ
- Volatilité annualisée de l'actif risqué : σ
- Taux annualisé pour l'actualisation du passif : t_a (se déduit de t_m par les formules de conversion usuelles)

Le taux sans risque r est fixé à 1,5 %.
Les niveaux de performance et de volatilité de l'indice actions peuvent être calibrés à partir des données historiques (suivant la zone géographique considérée)

A ce stade de l'étude, nous proposons les choix suivants : μ = 5,5 % et σ = 20 %

Enfin, pour le taux annuel d'actualisation de passif, nous proposons de retenir pour le moment le taux sans risque r. Ce choix permet (au vu des résultats de simulations) de garantir une probabilité négligeable de surplus final négatif. Une autre approche (moins conservatrice) pourrait cependant consister à utiliser une moyenne pondérée entre le taux sans risque et la performance annuelle de l'indice action, suivant l'allocation initiale de l'actif du portefeuille :

$$t_a = w_a \times \mu + (1 - w_a) \times r$$

Où w_a est le pourcentage du portefeuille initialement investi dans l'actif risqué.

I-2-2-5 *Calibrage des paramètres de gestion*

Entre chaque date de versement de flux, l'allocation d'actif du fonds est mensuellement ajustée suivant une stratégie à coussin de type CPPI. Comme vu précédemment, cela consiste à remettre, à chaque date mensuelle t, le montant $E(t)$ placé en actif risqué égal à un multiple m du surplus du portefeuille (également appelé « coussin ») à cette même date. Cette opération d'ajustement de l'exposition se fait (en théorie) de façon instantanée à la date t, sans que le marché n'ait eu le temps de fluctuer dans l'intervalle.

Dans la gestion CCPI « classique » nous avons ainsi :

$$E(t) = m \times C(t)$$

m est le « multiplicateur » associé à la gestion CPPI
$C(t)$ est le « coussin » à la date t : $C(t)$ = Actif du portefeuille (t) –Passif (t) = surplus (t)

Une contrainte supplémentaire a été ajoutée par rapport à la gestion CPPI classique, elle découle du fait que l'actif du portefeuille (et également le coussin) est susceptible d'être augmenté à fréquence annuelle par le versement du flux.

Plus précisément la règle de gestion est la suivante : $E(t) = Min(m \times C(t), A(t))$ où $A(t)$ est la valeur de l'actif du portefeuille à la date mensuelle *t*. Cette contrainte traduit l'impossibilité d'investir en actions un montant supérieur à l'actif du portefeuille. Nous pouvons également noter que les versements annuels des souscripteurs, ayant pour effet d'augmenter le coussin, permettent de réduire le risque que l'actif soit monétarisé en cas de chute brutale des marchés.

Les paramètres spécifiques à la gestion qui doivent être calibrés sont donc les suivants :

- Le multiplicateur *m*
- La fréquence entre deux réajustements de l'exposition : *t*

Pour cette étude nous nous sommes limités, à titre illustratif, au cas d'un rebalancement mensuel de l'exposition. Une fréquence hebdomadaire, voire quotidienne, permet d'améliorer la gestion de la garantie en tenant compte de la possibilité de « décrochement » brutal du marché entre deux dates de réajustement. Cependant, en pratique, un compromis devra être trouvé entre la précision de la gestion du coussin et la limitation des coûts de transactions (ces coûts sont réduits si les opérations d'achat/vente se font via des contrats futures sur l'indice action de référence).

Le choix du multiplicateur *m* se fait ici au moyen de simulations Monte-Carlo de la gestion mensuelle du fonds :

- Pour différentes valeurs de *m* (en l'occurrence de 0,5 à 20 avec un incrément de 0,1), nous simulons 1000 trajectoires du portefeuille (actif et passif) en intégrant l'échéancier des flux pour la tranche d'âge considérée.
    - Nota : nous nous sommes limités à ce stade à la tranche d'âge des 25 ans en 2010.

- Plusieurs critères d'évaluation des résultats en fonction du multiplicateur m ont été envisagés, ils se fondent en particulier sur le surplus ainsi que sur la probabilité de perte (à l'échéance ou en cours de vie du fonds).
- La liste de ces différents critères est la suivante :
    - Ratio de Sharpe[15] du surplus sur la trajectoire
    - Ratio de Sharpe du surplus final
    - VaR à 95% du surplus final
    - Probabilité de perte (surplus négatif) sur une trajectoire
    - Richesse finale moyenne
    - VaR à 95 % de la richesse finale
    - Surplus final moyen
    - Probabilité de surplus final négatif

I-2-2-6 *Résultats*

Nous présentons les graphiques des niveaux des critères fondés sur le surplus, en fonction du choix de *m* (1 000 simulations en retenant l'échéancier de la tranche 25 ans en 2010).

- Ratio de Sharpe du surplus final
- VaR à 95 % du surplus final

[15] Rapport entre la moyenne des surplus obtenus sur les différentes trajectoires à la date finale et leur écart-type à la même date.

- Probabilité de surplus négatif sur les trajectoires (en cours du vie du fonds)
- Surplus final moyen
- Probabilité de surplus final négatif

Le choix de $\underline{m=2,5}$ apparaît comme un bon compromis entre la réduction du risque et la performance offerte aux investisseurs : d'une part les indicateurs de performance (ratio de Sharpe du surplus final, surplus final moyen) se stabilisent à partir de la valeur m=2,5 et d'autre part les probabilités d'un surplus négatif (sur les trajectoires ou au final) sont quasi nulles à partir de la même valeur (cf. graphique 64, 65, 66, 67, 68, 69 et 70).

A titre illustratif, nous indiquons dans les graphiques ci-dessous, pour cette valeur optimale du multiplicateur, l'allure de la trajectoire moyenne correspondante de l'actif du portefeuille et de la garantie (cf. graphique 69) ainsi que les 1000 trajectoires simulées du surplus (cf. graphique 70). Les graphiques 67 et 68 montrent quant à eux que la probabilité d'avoir un surplus négatif augmente (que ce soit en cours de vie du fonds ou à l'échéance) suite à l'augmentation de m . Cela est tout à fait conforme à nos attentes puisque m traduit le degré d'exposition à l'actif risqué et donc au risque assumé par l'investisseur. A ce niveau, notons que les irrégularités qui apparaissent sur certaines courbes au delà de m=2,5 (cas du graphique 66 par exemple) sont expliquées par le fait que le nombre de simulations effectuées (1 000) n'est pas encore assez élevé pour faire disparaître cette irrégularité autour de la tendance.

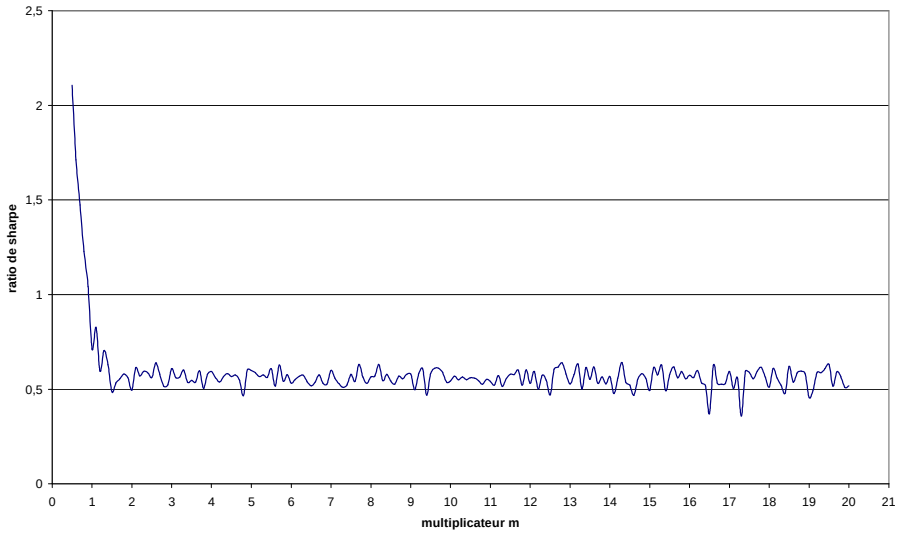


Fig. 64 : *Ratio de Sharpe du surplus final en fonction de m*

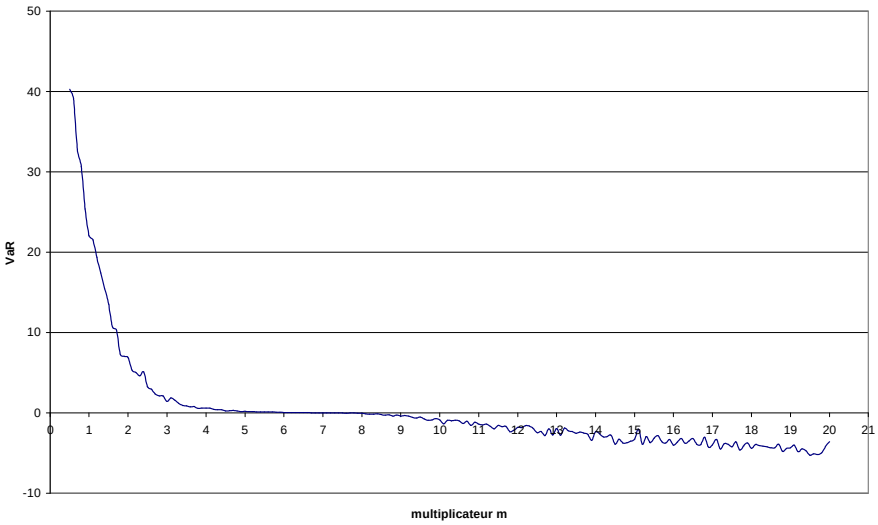


Fig. 65 : *VaR à 95 % du surplus final en fonction de m*

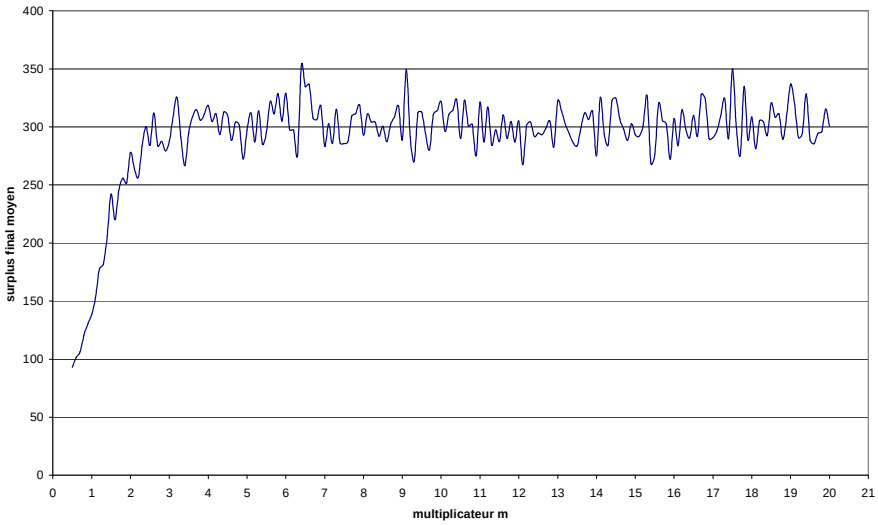


Fig. 66 : *Surplus final moyen en fonction de m*

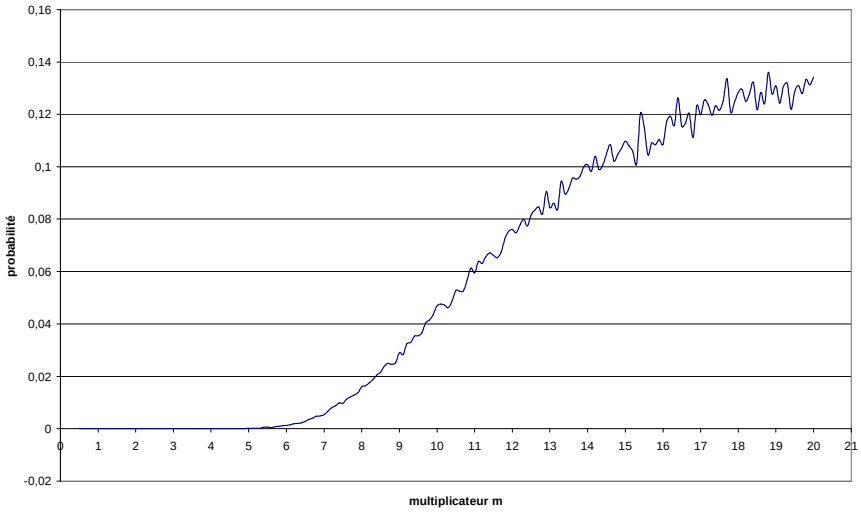


Fig. 67 : *Probabilité de surplus négatif en cours de vie du fonds (en fonction de m)*

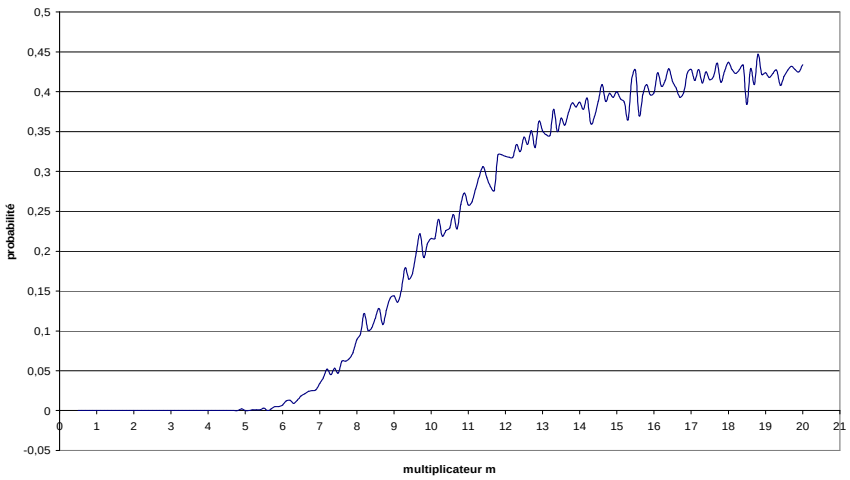


Fig. 68 : *Probabilité de surplus final négatif en fonction de m*

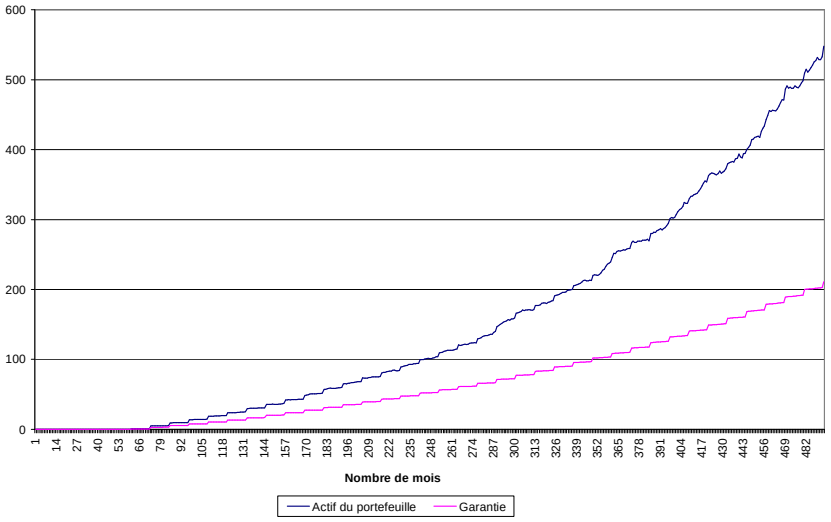


Fig. 69 : *Trajectoire moyenne du portefeuille (actif et garantie) pour m=2.5*

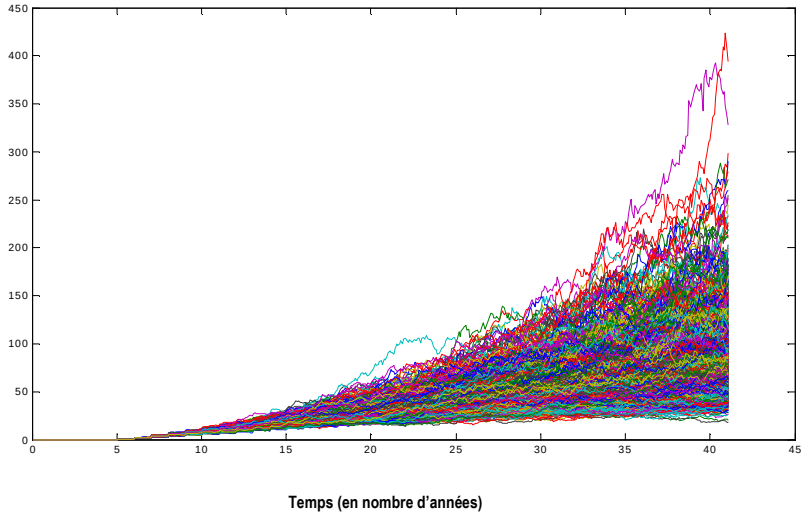


Fig.70 : *Allure des 1000 trajectoires simulées du surplus (m=2.5)*

La stratégie d'assurance de portefeuille reste une des stratégies les plus intéressantes en pratique. Un inconvénient majeur de cette approche consiste dans le fait qu'elle limite le champ du choix de la règle de gestion pour l'investisseur : elle réduit sa problématique à l'unique gestion de la proportion constante (cf. Hilli et al. [2007]). Cet éventuel manque de flexibilité, en termes de choix des éléments du système d'optimisation (que ce soit au niveau de la fonction objectif ou au niveau des contraintes), nous mènent à explorer d'autres techniques pouvant traiter des cas plus généraux.

I-3 Programmation dynamique

Dés lors que l'on veut aborder les problèmes d'allocation inter-temporelle d'actifs financiers, nous faisons souvent recours à des méthodes plus sophistiquées. S'offre également la possibilité de se placer soit en temps discret soit en temps continu. Le cas multi-périodique s'avère *a priori* très lourd à manipuler. En temps continu, la résolution de ce type de problème peut par exemple se faire via les techniques de programmation dynamique : soit à partir d'une approche martingale (Cox et Huang [1989]) soit à partir du contrôle stochastique (Merton [1971], Bellman [1957]).

Au niveau de la littérature, nous pouvons citer les travaux de Hainaut et al. [2007], qui analysent la politique de dividende et l'allocation d'actif pour un fonds de pension. Ils considèrent un marché financier composé de trois actifs : monétaire (ou *cash*), actions et obligations glissantes (*rolling bond*). Les taux d'intérêt sont supposés suivre un modèle de Vasicek [1977] alors que la mortalité de la population assurée est modélisée par un processus de Poisson. Ils déterminent les politiques d'investissement et de dividende qui maximisent l'utilité des dividendes et du surplus final sous une contrainte budgétaire. En particulier, les solutions sont développées pour les fonctions d'utilité de type CRRA[16]. La méthodologie est basée à la fois sur l'approche Cox et Huang [1989] et sur le principe de la programmation dynamique (Merton [1971]).

Hainaut et al. [2005] étudient le problème de l'allocation optimale d'actif en temps continu pour un fonds de pension sous la contrainte de la VaR. La mortalité est modélisée par un brownien géométrique et les flux payés aux adhérents sont considérés comme déterministes. Le marché financier est composé d'actifs risqués suivant des mouvements browniens géométriques. Nous considérons le cas d'un gérant de fonds qui cherche à maximiser l'utilité quadratique du ratio de solvabilité, définie comme la valeur de marché des actifs divisée par les réserves mathématiques, sous une contrainte de type Valeur à Risque (VaR). La méthode du multiplicateur de Lagrange (Föllmer et al. [1998]) est combinée avec l'équation de Hamilton-Jacobi-Bellman HJB (Merton [1971]) afin d'insérer la contrainte de gestion des risques dans le cadre de la solution. Une méthode numérique est ensuite développée afin d'obtenir une approximation de la solution de l'équation HJB. Finalement, est analysé l'impact du choix de la contrainte VaR sur la stratégie d'investissement optimale.

Rudolf et al. [2004] présentent un modèle de sélection de portefeuille dans un cadre inter-temporel dynamique aussi. L'objectif est de maximiser l'espérance d'utilité du surplus des actifs nets du passif. A partir de Merton [1971], il est supposé que les rendements des actifs et du passif suivent des processus d'Itô. Ces deux rendements vont permettre de définir des

[16] Pour « *Constant Relative Risk Aversion* ». L'expression analytique générale pour cette classe de fonction d'utilité est : $U(R)=\frac{R^{1-\beta}}{(1-\beta)}$ où β est le coefficient d'aversion pour le risque (cf. Campbell et al. [2001]).

variables d'état (tel que le surplus, ratio de financement, etc.). L'optimalité est obtenue pour un investisseur détenant quatre fonds : le portefeuille de marché, le portefeuille de couverture de la variable d'état, le portefeuille de couverture des passifs et enfin l'actif sans risque. A la différence du résultat trouvé par Merton, dans le cadre d'analyse dynamique actif-unique (et non pas actif-passif), la couverture du passif est indépendante des préférences et dépend seulement du ratio de financement. Avec une fonction d'utilité de type HARA [17], l'investissement dans le fonds de couverture de la variable d'état est indépendant des préférences de l'individu. Enfin, avec des fonctions d'utilité logarithmiques, l'investissement en portefeuille de marché dépend seulement du ratio de financement courant. Une étude empirique prenant le cas d'un investisseur américain avec une possibilité de diversification internationale est menée.

Yen et al. [2003] étudient l'allocation d'actif optimale et dynamique pour un fonds de pension dans un contexte d'opportunités d'investissement qui évoluent dans le temps. Les sources de changement dans ce modèle sont dues au changement de la volatilité et des primes de risque au cours du temps. Ils étudient l'allocation d'actif pour un fonds de pension, en cas de volatilité indexée sur le temps, en utilisant l'approche de passif de Sharpe et Tint [1990] ainsi que la méthodologie inter-temporelle d'évaluation des actifs de Merton [1971] concernant les stratégies d'investissement des fondations universitaires (*university endowment funds*). Ils développent un modèle inter-temporel et déduisent une solution explicite en utilisant des méthodes de perturbation. A la différence de Merton [1971], ils proposent une nouvelle composante de couverture du passif dans le cadre de l'allocation dynamique, en plus de la composante de couverture inter-temporelle des fonds de pension. Cela permet de mieux couvrir le changement du passif au cours du temps.

Dans ce qui suit, nous focaliserons notre intérêt sur les techniques de contrôle stochastique optimal (cf. Merton [1971]). En particulier, une résolution des problèmes d'allocation d'actifs dans le cadre multi-périodique peut être obtenue à l'aide de l'équation d'Hamilton Jacobi Bellman (HJB). Nous nous référons, dans ce qui suit, au travail de Malliaris et al. [1982] pour développer l'équation de HJB dans le cadre de la théorie de la commande ainsi que pour une illustration de ces techniques dans le cas de la maximisation de l'utilité de la richesse finale de l'investisseur.

<u>L'équation de HJB dans le cadre de la théorie de la commande :</u>

Nous considérons que l'évolution stochastique de n variables x, qui peut être éventuellement contrôlée par une commande V à la disposition d'un centre de décision, obéit à l'équation différentielle stochastique vectorielle suivante :

$$dx = \phi(V, x, t)dt + \sigma(V, x, t)dz \qquad (1)$$

[17] Pour « *Hyperbolic Absolute Risk Aversion* » (cf. Zenios [2007]). Une fonction d'utilité $U(R)$ est une fonction d'utilité HARA si elle s'écrit : $U(R) = \frac{1-\gamma}{\gamma}\left(\frac{\beta R}{1-\gamma} + \eta\right)^{\gamma}\beta$ où β, $\gamma \neq 1$ et η sont des paramètres réels (restrictions si $\gamma > 1$ pour le domaine de validité de la fonction d'utilité). Les utilités quadratiques, CRRA (*Constant Relative Risk Aversion*) et CARA (*Constant Absolute Risk Aversion*) sont des cas particuliers de HARA.

$$x(0) = x_0$$

Avec :
- z est un mouvement brownien de dimension d
- V est le vecteur de commande de dimension k et supposé se trouver dans un domaine U.
- ϕ est un vecteur (n,1) donnant la dérive de x.
- σ est la matrice (n, d) de diffusion de x, nous désignons $C = \sigma\sigma'$.

Soit D un domaine de $[0,T]\times R^n$, tel que lorsque (t,x) atteint la frontière ∂D de ce domaine, le système s'arrête.

Nous définissons le temps d'arrêt τ par : $\tau = inf\{t \geq 0 / (t,x(t)) \in \partial D\} \wedge T$.

Le système doit être contrôlé de telle façon que la commande V permette de maximiser le critère :

$$E\left[\int_0^\tau u(V,x,t)dt + \chi(x,\tau)\right] \qquad (2)$$

u et χ sont des fonctions numériques. Dans le contexte économique, ce sont des fonctions d'utilité. Nous définissons l'opérateur L_V par :

$$L_V = \sum_{i=1}^{n} \phi_i(V,x,t)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} C_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}$$

Nous considérons la fonction $K(V,x,t) = E_t\left[\int_t^\tau u(V,x,s)ds + \chi(x,\tau)\right]$ ainsi que la fonction $J(x,t) = \max_{V\in U} K(V,x,t)$.

Lorsque la frontière est touchée, le système s'arrête et fournit au centre de décision la performance $\chi(x,\tau)$. Un problème de commande optimale consiste à trouver la commande V qui maximise le critère (2) lorsque le système évolue selon l'équation (1) ci-dessus, le théorème suivant (équation de HJB) apporte une solution à ce problème :

Si J est de classe C2, et s'il existe une commande optimale V^0, alors : à chaque instant le contrôle optimal V^* vérifie : $u(V^*,x,t) + L_{V^*}K(V,x,t) \geq u(V,x,t) + L_V K(V,x,t) \quad \forall V \in U$

La fonction J, vérifiant cette équation, satisfait également l'équation aux dérivées partielles (EDP) de premier ordre appelée équation d'Hamilton-Jacobi-Bellman :

$$J_t + \max_{U\in V}[L_V J + u] = 0 \qquad \forall (x,t) \in D$$
$$J(x,t) = \chi(x,t) \qquad \forall (x,t) \in \partial D$$

La méthode à utiliser consiste, dans un premier temps, à trouver la commande V^* en fonction de J, telle qu'obtenue par la maximisation de $L_V J + u$, puis dans un deuxième temps à la reporter dans l'équation de HJB, que nous cherchons à résoudre. Celle-ci résolue nous obtenons alors la solution définitive. C'est un peu la même démarche qu'avec les

multiplicateurs de Lagrange, mais en plus compliquée. En fait cette approche est plus théorique que pratique car l'équation de HJB est très difficile à résoudre et conduit assez rarement à des solutions explicites, hormis dans quelques cas particuliers.

Illustration dans le cas de la maximisation de l'utilité de la richesse finale de l'investisseur

Pour illustrer l'utilisation de HJB, nous considérons le cas simple de choix de portefeuille lorsque l'agent cherche à maximiser uniquement l'utilité de sa richesse finale. Nous supposons que la fonction d'utilité est de type CRRA (Constant Relative Risk Aversion) :

$$\chi(W(T)) = \frac{W(T)^{\gamma}}{\gamma}$$

où $\gamma < 1$ et $\gamma \neq 0$.

Le système comporte ici une seule variable d'état pour le système : la richesse qui suit l'équation $dW = [w(\mu - r)W + rW]dt + w\sigma W dz$. La commande est donnée par une seule variable non contrainte : w, la part à investir dans l'actif risqué. L'équation HJB s'écrit dans ce cas comme suit :

$$J_t + \max_w \left[J_W \left(w(\mu - r)W + rW \right) + \frac{1}{2} w^2 \sigma^2 W^2 J_{WW} \right] = 0$$

La maximisation conduit à obtenir la commande w en fonction de J, sous la forme :

$$w^* = -\frac{J_W}{W J_{WW}} \frac{\mu - r}{\sigma^2}$$

En reportant cette valeur dans l'équation de HJB, il vient :

$$-\frac{1}{2} \frac{J_W^{\,2}}{J_{WW}} \left(\frac{\mu - r}{\sigma} \right)^2 + rWJ_W + J_t = 0 \qquad (3)$$

avec la condition limite : $J(W(T),T) = \frac{W^{\gamma}(T)}{\gamma}$ et $J(0,t) = 0$ pour tout t.

La résolution de l'équation (3) se fait en cherchant une solution du type : $J(W,t) = h(t)\frac{W^{\gamma}}{\gamma}$.

Nous obtenons alors :

$$h' \frac{W^{\gamma}}{\gamma} + rhW^{\gamma} + \frac{1}{2} h \frac{W^{\gamma}}{1-\gamma} \left(\frac{\mu - r}{\sigma} \right)^2 = 0$$

Avec la condition limite $h(T) = 1$.

Soit encore :

$$-\frac{h'}{h} = \gamma \left(r + \frac{1}{2(1-\gamma)} \left(\frac{\mu - r}{\sigma} \right)^2 \right)$$

En posant :

$$\eta = \gamma \left(r + \frac{1}{2(1-\gamma)} \left(\frac{\mu - r}{\sigma} \right)^2 \right)$$

nous pouvons écrire : $-\frac{dh}{h} = \eta dt$ avec $h(T) = 1$.

Après intégration de t à T, nous obtenons : $J(W,t) = e^{\eta(T-t)} \frac{W^{\gamma}}{\gamma}$.

La proportion à investir dans l'actif risqué est donc : $w^* = \frac{\mu - r}{\delta \sigma^2}$ avec $\delta = 1 - \gamma$.

Cette part est d'autant plus importante que l'écart entre le rendement attendu de l'actif risqué et le taux sans risque est grand et d'autant moins que la volatilité de l'actif risqué ainsi que l'aversion relative au risque δ de l'agent sont grandes. L'objectif de cet exemple est de montrer de façon détaillée comment fonctionne la recherche d'une solution optimale à l'aide de l'équation HJB.

Plus le nombre de paramètres dans le modèle augmente plus l'obtention d'une solution analytique au problème d'optimisation dynamique devient difficile et souvent impossible (cf. Brandimarte [2006]). Le recours au Programmation Stochastique (PS) permet la prise en compte de modèle plus proche de la réalité. Ces techniques constituent l'objet de l'étude détaillée de la section suivante.

II- L'allocation stratégique d'actifs dans le cadre de la programmation stochastique

II-1 Introduction

Dans le cadre général de la résolution des problèmes d'optimisation stochastique deux approches inspirées des techniques de recherche opérationnelle sont souvent présentées dans la littérature : la programmation dynamique (*Dynamic Programming*) et la programmation stochastique (*Stochastic Programming*). Ces deux approches ont différents points communs. Alors que la programmation dynamique est quant à elle souvent utilisée dans le monde de l'économie (en particulier en théorie de la consommation et la théorie des options), la programmation stochastique est souvent appliquée par la communauté des ingénieurs en particulier dans le domaine de la planification de la production.

La programmation dynamique (en particulier le contrôle optimal) est un outil puissant et performant dans le domaine de la finance pour les problèmes relativement simples et permet d'obtenir des résultats plausibles de point de vue théorique et pratique. L'inconvénient majeur de ce type de modèle est qu'il ne permet pas de traiter des systèmes d'optimisation complexes et réalistes, principalement pour des raisons d'implémentation numérique (cf. Brandimarte [2006]). L'addition de contraintes est souvent une sérieuse entrave à la technique de programmation dynamique puisqu'elle augmente la dimension de l'espace des variables d'état, ce qui rend difficile la résolution du problème. La nécessité de passage par des hypothèses simplificatrices met ainsi en cause la fiabilité des résultats obtenus, c'est le cas de la gestion actif-passif d'un fonds de retraite ou d'un fonds de pension.

Les modèles de programmation stochastique, avec leurs outils de traitement mathématique poussés (analyse de convexité, théorie de dualité, etc.), sont par contre plus aptes à faire face à des problèmes complexes et à intégrer une multitude de variables et de contraintes. De façon générale, la programmation stochastique (PS) est le nom générique des modèles d'optimisation qui sont utilisés dans le domaine de la planification. Leur application par les ingénieurs dans le domaine de la recherche opérationnelle est répandue : allant de la gestion

de l'alimentation en électricité jusqu'à la planification des réseaux de télécommunication (cf. Dantzig et al. [1990], Escudero et al. [1993]).

Plusieurs études montrent que les techniques de PS surperforment les techniques décrites précédemment, en particulier la programmation dynamique et l'optimisation basée sur la stratégie *Fixed-Mix* (cf. Cariño et Turner [1998], Ziemba [2003]). Nous nous proposons d'étudier les modalités de leur adaptation dans le cadre d'une étude de gestion actif-passif. En particulier, nous focalisons notre intérêt sur les modèles de programmation stochastique avec recours (cf. Dantzig et al. [1990], Escudero et al. [1993], Birge et Louveaux [1997]).

Ainsi, nous nous penchons dans cette section sur une l'introduction d'une approche innovante fondée sur les techniques de programmation stochastique. Dans un premier temps, nous explorons le cadre d'utilisation de ces techniques en s'appuyant sur un exemple pratique dans le domaine de la planification de la production. Le lien avec le monde de la finance est mis en évidence au terme de cette partie. Ensuite, nous formulons, de point de vue mathématique, le problème de programmation stochastique avec recours et nous recensons certains outils de résolution de ce problème. La mise en place effective d'un modèle d'ALM dynamique basé sur les techniques de PS fera quant à lui l'objet de la section suivante (section III).

II-2 Présentation de la programmation stochastique

Le comité de programmation stochastique COSP (*Committee on Stochastic Programming*[18]) définit la programmation stochastique comme « un cadre de modélisation des problèmes d'optimisation incluant l'incertitude ». Alors que les problèmes d'optimisation déterministes sont formulés avec des paramètres et des coefficients connus, les problèmes du monde réel incluent souvent des paramètres aléatoires. La programmation stochastique cherche à présenter la solution optimale en se basant sur le fait que la distribution de probabilité des données est connue ou qu'elle peut être estimée.

La programmation stochastique est un programme mathématique dans lequel certains des paramètres du problème étudié sont aléatoires; la programmation stochastique linéaire est le cas particulier le plus simple de la PS. Parmi les hypothèses nécessaires pour l'utilisation de ces techniques nous pouvons citer les deux suivantes : d'une part, la distribution de probabilité de l'incertitude (sur les paramètres) est soit connue soit pouvant être estimée et d'autre part les probabilités sont indépendantes des décisions à prendre (cf. Zenios [2007]).

Les concepts basiques des modèles d'ALM sous l'incertitude et avec PS ont été développés par Kallberg, White et Ziemba [1982] ainsi que Kusy et Ziemba [1986]. Plusieurs développements ont été ensuite présentés. Parmi les modèles les plus adoptés en pratique, nous trouvons le modèle de Russel-Yasuda Kasai développé par Cariño et al. [1994]. Il s'agit d'un modèle d'ALM pour une compagnie d'assurance japonaise Yasuda basé sur la programmation stochastique multi-périodique (*multistage stochastic programming*). Selon ces auteurs, le modèle permet à Yasuda de combiner à la fois les outils sophistiqués utilisés en matière de prise de décision et les outils de gestion des risques. Ceci permettra d'avoir une analyse plus fiable et plus performante du positionnement de la société vis-à-vis des différents risques auxquels elle est exposée.

[18] Un groupe international de chercheurs qui développent des modèles, des méthodes et des théories pour la prise de décision dans le contexte d'incertitude : http://stoprog.org/

Un autre modèle d'ALM dynamique basé sur la programmation stochastique appelé « Gestion actif-passif assistée par ordinateur » (en anglais « *Computer-aided asset/liability management* ou CALM ») est présenté par Consigli & Dempster [1998] : ces derniers montrent que le modèle CALM est construit en tenant compte de l'incertitude affectant à la fois les actifs (composant le portefeuille ou existant sur le marché) et les engagements (sous forme de scénarios dépendant de flux de décaissement et de coûts des opérations d'emprunt).

Le modèle d'ALM dynamique présenté par Dert [1995] reste également un modèle de référence en matière d'exploitation des techniques de PS pour le pilotage technique d'un fonds de pension. Il est appliqué dans le cas des fonds de pension à prestation définie. Dert présente un modèle d'optimisation qui étudie conjointement la politique d'investissement et la politique de financement d'un fonds de pension à prestations définies en tenant compte de l'incertitude économique affectant ses engagements futurs vis-à-vis de ses adhérents.

Ce modèle permet aussi de déterminer les stratégies d'ALM à adopter, que ce soit au niveau des décisions d'investissement ou au niveau des taux de cotisation dans un contexte d'incertitude. Dert constate que les décisions sont différentes que nous utilisons les stratégies d'ALM dynamiques ou des stratégies d'ALM statiques. De même, en utilisant un modèle dynamique d'ALM, il a constaté que les stratégies proposées sont à moindre coût de financement et que les probabilités de sous financement sont significativement plus faibles.

La programmation stochastique traite en général des problématiques de planification en présence d'incertitude. Ce type de problème se rencontre le plus souvent dans les modèles qui à la fois s'échelonnent dans le temps et où les décisions prises à une période donnée ont des implications sur celles prises à la période suivante (et évidemment où l'aléa continue à exister au niveau des paramètres des périodes subséquentes). Ce type de modèle est souvent formulé, résolu (de façon analytique ou numérique) et analysé dans le but de fournir une information utile au preneur de décision.

Selon Kaut et Wallace [2003], la programmation stochastique a eu une popularité croissante dans la communauté des mathématiciens programmeurs. La puissance actuelle des outils informatiques permet aux utilisateurs d'ajouter les aspects stochastiques aux modèles déterministes difficiles à résoudre il y a quelques années. Dans ce cadre, la programmation stochastique peut être vue comme des modèles de programmation déterministe avec incertitude au niveau de certains paramètres. Au lieu d'avoir des valeurs uniques, ces paramètres sont décrits par des distributions (dans le cas d'une seule période), ou par des processus stochastiques (dans le cas multi-périodique).

Les modèles de programmation stochastique les plus appliqués et étudiés sont les programmes avec recours (*with recourse,* cf. Brandimarte [2006]). Dans le cas mon-périodique par exemple, la démarche générale de ce type de modèle suppose la prise de deux décisions : la première au début de la période et donc avant la réalisation de l'évènement aléatoire et la deuxième à la fin de la période et après la réalisation de l'évènement aléatoire. Autrement dit, le planificateur aura à prendre une décision dans la première étape compte tenu de l'information disponible à cette date, avant qu'un événement aléatoire ait lieu. Une décision de recours peut ensuite être prise dans une deuxième étape afin de compenser les éventuelles conséquences négatives dues à une mauvaise décision initiale.

La politique optimale dans ce type de modèle permet de définir les décisions à l'étape initiale ainsi que les différents recours ultérieurs correspondant à chaque réalisation possible des

éléments aléatoires. Ces résultats peuvent ensuite être étendus pour le cas de la programmation stochastique multi-périodique. La partie suivante illustre de façon détaillée les principes de la programmation stochastique « avec recours » dans le cadre d'un problème de planification de la production. Le même principe sera par la suite étendu au cadre financier.

II-3 Illustration de la programmation stochastique « avec recours » dans le cas de la planification de la production

Les problèmes rencontrés dans le domaine de la planification de la production constituent le contexte classique d'utilisation des techniques de programmation stochastique (cf. Escudero et al. [1993]).

II-3-1 *Cas mono-périodique (à deux étapes)*

La programmation stochastique avec recours dans le cas d'une seule période passe par deux étapes :

- dans la première étape, qui correspond au début de la période, il y aura une prise de décision sur un élément incertain.
- dans la deuxième étape, qui correspond à la fin de la période, et après la réalisation de la variable aléatoire, des décisions supplémentaires peuvent être prises afin d'éviter que les contraintes soient non respectées.

Autrement dit, dans la seconde étape nous avons **recours** à un degré complémentaire de flexibilité afin de préserver la faisabilité (avec un coût relatif bien évidemment). Notons en particulier, qu'à cette deuxième étape les décisions qu'on prend dépendront des réalisations observées des éléments stochastiques.

Nous illustrons ce simple modèle (avec recours et à deux étapes) dans le cadre de la mise en place d'un plan de production[19]. Supposons que nous ayons à décider de la quantité x_1 du produit X à produire. Chaque unité de X que nous produisons coûte 2 €. La quantité x_1 est supposée répondre à une demande des consommateurs déterminée à la fin de la période. La demande est donc stochastique avec une distribution de probabilité discrète : elle est égale à D_s avec une probabilité p_s ($s = 1,...,S$) et ainsi S demandes futures possibles.

Il est également supposé que la demande des consommateurs devra être satisfaite quel que soit la décision sur la quantité initialement produite (x_1). En cas d'une demande effective supérieure à cette quantité, une solution sera par exemple d'acheter la quantité manquante du produit X auprès d'un fournisseur externe mais avec un coût unitaire de 3 € (c'est à dire avec **recours** à une source additionnelle de l'offre avec un coût correspondant : coût du recours). La question posée est donc quelle quantité optimale x_1^* devons nous produire avant de connaître la réelle demande des consommateurs ?

Ci-dessous, nous illustrons notre exemple par des chiffres avec: $S = 2$ et $D_1 = 500$, $p_1 = 0{,}6$; $D_2 = 700$, $p_2 = 0{,}4$.

[19] en s'inspirant de Beasley [1990] et précisément de http://people.brunel.ac.uk/~mastjjb/jeb/info.html

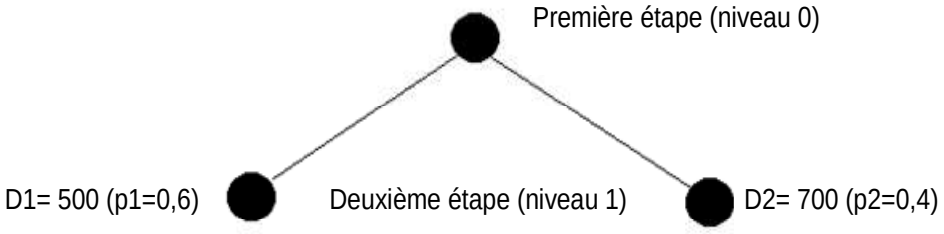


Fig. 71 : *Illustration du problème de planification à deux étapes (dans le cas d'une seule période)*

Nous avons donc à décider de la quantité à produire à la date actuelle, avant que la demande ne soit connue. Si nous avons à produire 600 unités par exemple et que la demande effective sera de 500 nous n'aurons pas de problème. Cependant, si la demande est de 700 nous aurons besoin de recourir à l'achat de 100 unités supplémentaires.

Nous procédons de façon classique à la modélisation du problème, en définissant les variables. Soit $x_1 >= 0$ le nombre d'unités à produire à l'instant de départ (à la première étape). Etant donné S scénarios, nous associons un scénario à chaque variable de recours à la deuxième étape, soit donc $y_{2s} >= 0$ le nombre d'unités de X à acheter auprès de notre fournisseur externe à la seconde étape pour le scénario s, lorsque la réalisation de la demande est D_S ($s = 1,...,S$).

Ainsi, les contraintes qui garantissent la satisfaction de la demande seront :
$x_1 + y_{2s} >= D_s$ Avec $s = 1,...,S$

Notons qu'il faut avoir ici l'égalité non stricte ($>=$) car la quantité x_1 à produire pourra dépasser la demande effective des consommateurs. Si nous retenons $x_1 + y_{2s} = D_s$ avec l'hypothèse $y_{2s} >= 0$, $\forall s = 1,...,S$, nous seront emmenés à supposer que : $x_1 <= min[D_s / s = 1,...,S]$ ce qui limite l'univers de décision.

Pour la fonction objectif, nous avons un coût de $2x_1$ supporté avec certitude et S coûts possibles coûtant chacun $3y_{2s}$ et se réalisant chacun avec une probabilité p_s. Notons qu'en pratique, un seul scénario sera observé effectivement mais qu'avant sa réalisation nous nous disposons que de sa distribution de probabilité. Dans notre exemple, nous avons choisi intuitivement la fonction objectif comme étant la minimisation du coût total espéré qui est égal à : $2x_1 + \sum_{s=1}^{S} p_s(3y_{2s})$

Notre programme stochastique avec recours est :

$$\textit{Minimiser } 2x_1 + \sum_{s=1}^{S} p_s(3y_{2s})$$
$$\text{S/c : } x_1 + y_{2s} >= D_s \qquad s = 1,...,S$$
$$x_1 >= 0$$
$$y_{2s} >= 0 \qquad s = 1,...,S$$

Notons qu'il s'agit là d'un programme déterministe. Nous pouvons aussi exiger que x_1 et y_{2s} soient des entiers.

La résolution de ce programme stochastique nous indique :

- une valeur de x_1 qui est la quantité de X que nous allons produire maintenant,
- l'ensemble des valeurs possibles de y_{2s} ,

Notons qu'une seule valeur y_{2s} parmi les S valeurs possibles sera importante lorsque la demande des consommateurs sera connue.
Pour résumer, dans un programme stochastique avec recours :

- nous prenons des décisions à la première étape, sachant seulement la distribution de la probabilité des éléments stochastiques,
- nous avons recours, avec des coûts relatifs, à la deuxième étape à des variables permettant d'assurer la faisabilité des contraintes, ces variables seront différentes pour différentes réalisations des éléments stochastiques,
- nous minimisons le coût total espéré: somme des coûts connus certains, issus des décisions de la première étape et des coûts espérés issus des décisions de la seconde étape.

Notons que le fait que l'élément stochastique (la demande des consommateurs) ait une distribution discrète a simplifié la formulation du problème. Si nous sommes en revanche dans le cas d'une distribution continue, alors le problème mathématique associé à la formulation et à la résolution deviendra plus réaliste et plus pertinent en pratique. L'objectif standard dans le cas de la programmation stochastique avec recours (à deux étapes) est de minimiser le coût à la première étape ainsi que le coût espéré à la deuxième étape (relatif à l'éventuelle décision de recours).

II-3-2 Cas multi-périodique (ou multi-étapes)

Une extension possible au modèle de programmation stochastique à deux étapes avec recours est de supposer des périodes supplémentaires.

Afin d'illustrer ce type de modèle, nous allons étendre le problème sur deux périodes : dans ce contexte nous nous projetons deux périodes dans le futur afin de planifier la production.
Nous considérons donc l'arbre binaire à trois étapes ci-dessous :

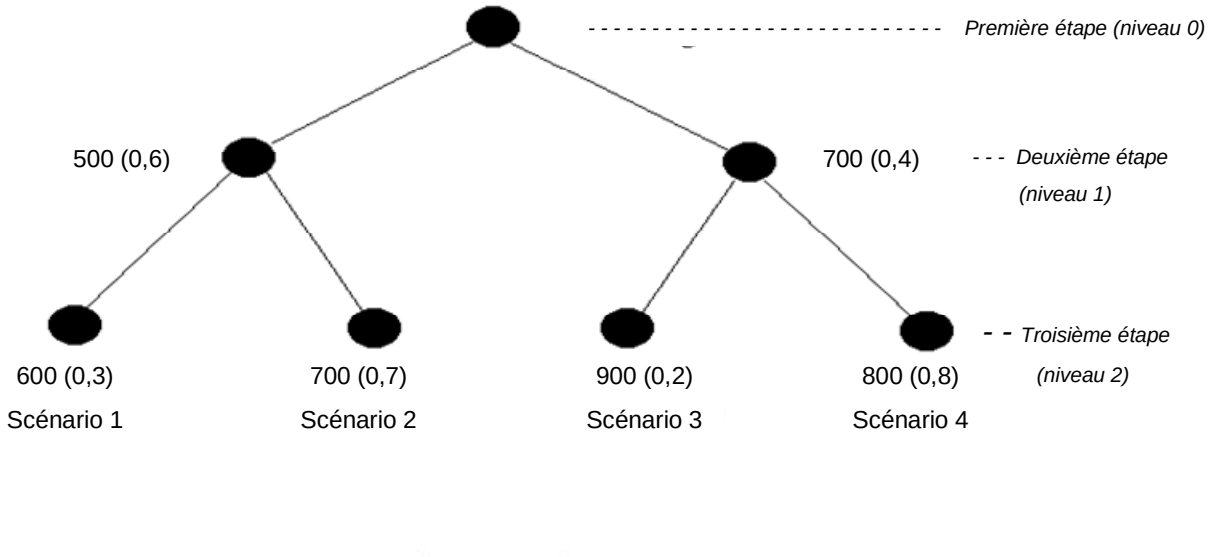


Fig.72 : *Illustration du problème de planification à trois étapes (dans le cas de deux périodes)*

A la première étape, nous prenons la décision de la quantité initiale de X à produire. A la deuxième étape, nous avons deux réalisations possibles de la demande stochastique :

- une demande de 500 avec une probabilité de 0,6
- une demande de 700 avec une probabilité de 0,4

Une fois que cette demande est réalisée et qu'une décision de recours a eu lieu à la deuxième étape, nous prenons la décision de la nouvelle quantité à produire, afin de satisfaire la demande de la troisième étape (fin de la deuxième période).

A la troisième étape, nous avons aussi deux réalisations possibles de la demande stochastique, mais celles-ci sont différentes selon la réalisation observée à la deuxième étape. Par exemple, si la demande réalisée à la seconde étape est de 500 alors les réalisations possibles à la troisième étape seront :

- une demande de 600 avec une probabilité de 0,3
- une demande de 700 avec une probabilité de 0,7

Notons ici qu'à chaque étape de l'arbre des scénarios, les probabilités correspondantes doivent avoir pour somme 1.

Cette arborescence à deux périodes représente $2^2=4$ scénarios possibles dans le futur :

Scénario	Deuxième étape	Troisième étape	Probabilité
1	500	600	0.6(0.3)=0,18
2	500	700	0.6(0.7)=0,42
3	700	900	0.4(0.2)=0,08
4	700	800	0.4(0.8)=0,32

Pour synthétiser, nous aurons le schéma décisionnel suivant en descendant l'arbre des scénarios :

- dans la première étape nous décidons de la quantité à produire,
- dans la deuxième étape nous aurons la réalisation des éléments stochastiques (la demande),

- une décision concernant le recours (achat ou non) est prise,
- une décision de la quantité à produire à la deuxième étape est prise,
- dans la troisième étape nous aurons la réalisation des éléments stochastiques (la demande),
- enfin, une décision concernant le recours (achat ou non) est prise à la troisième étape.

Ce schéma sera étendu bien évidemment si nous avons plus que trois étapes. Ce qui importe dans les modèles de programmation stochastique c'est la bonne identification des éléments stochastiques réalisés à une étape et ceux qui restent toujours inconnus.

Nous pouvons maintenant formuler notre programme stochastique à trois étapes. Soit donc :

$x_1 >= 0$ le nombre d'unités de X à produire à la date initiale (première étape)
$y_{2s} >= 0$ le nombre d'unités de X à acheter du fournisseur externe à la seconde étape pour le scénario s (s =1,…,4)
$x_{2s} >= 0$ le nombre d'unités de X à produire à la deuxième étape pour le scénario s (s =1,…,4)
$y_{3s} >= 0$ le nombre d'unités de X à acheter du fournisseur externe à la troisième étape pour le scénario s (s =1,…,4)

Notons que la décision de choix de la quantité à produire à la deuxième étape (x_{2s}) dépend du scénario réalisé sur la demande des consommateurs à cette étape (en particulier du niveau du stock disponible à la deuxième étape et de la demande future possible), ce type de décision est appelé « *scenario-dependent* ».

Pour des raisons de simplification, nous allons garder les mêmes coûts de production utilisé dans le modèle à deux étapes (2 € pour une unité produite, 3 € pour une unité achetée).

Le schéma décisionnel spécifié ci-dessus avec les variables considérées devient donc :

- dans la première étape : la décision de la quantité à produire donne x_1,
- dans la seconde étape : nous aurons la réalisation de l'élément stochastique (la demande),
- une décision concernant les variables de recours donne le y_{2s},
- dans la deuxième étape une décision de la quantité à produire : x_{2s},
- dans la troisième étape nous aurons la réalisation de l'élément stochastique (la demande),
- une décision concernant les variables de recours donne y_{3s}.

Considérant la première étape, les contraintes garantissant la satisfaction de la demande des consommateurs sont :
$x_1 + y_{2s} >= 500 \qquad s = 1,2$
$x_1 + y_{2s} >= 700 \qquad s = 3,4$

A la deuxième étape, nous pouvons avoir un stock d'unités provenant d'un éventuel excès de la production x_1 par rapport à la demande manifestée à cette étape ; nous appellerons cet excès l'inventaire et il servira à répondre à la demande future. Son niveau sera de :

$x_1 + y_{2s} - 500 \qquad s = 1,2$
$x_1 + y_{2s} - 700 \qquad s = 3,4$

Pour garantir la satisfaction de la demande à la troisième étape, nous avons : inventaire+ quantité produite+achat externe >= la demande :

$x_1 + y_{2s} - 500 + x_{2s} + y_{3s} >= 600 \qquad s = 1$
$x_1 + y_{2s} - 500 + x_{2s} + y_{3s} >= 700 \qquad s = 2$
$x_1 + y_{2s} - 700 + x_{2s} + y_{3s} >= 900 \qquad s = 3$
$x_1 + y_{2s} - 700 + x_{2s} + y_{3s} >= 800 \qquad s = 4$

En plus de ces contraintes, nous aurons besoin d'ajouter certaines autres contraintes, afin de garder la cohérence de notre démarche. En fait, dans la pratique, les valeurs de y_{21} et de y_{22} devront être les mêmes ainsi que les valeurs de y_{23} et de y_{24}. La raison est que le scénario 1 et 2 ont « un historique commun » jusqu'à la deuxième étape (pareil pour les scénarios 3 et 4). Nous pouvons donc ne pas différencier les variables de recours y_{21} et y_{22} car, à ce stade de projection, ne nous pouvons pas distinguer entre ces deux scénarios.

Nous ajoutons ainsi la contrainte de $y_{21} = y_{22}$ et de $y_{23} = y_{24}$. Ce type de contrainte est appelé « contrainte non anticipative » (*nonanticipativity constraint*), impliquant que nous ne pouvons pas anticiper le futur. Le même raisonnement conduit à déduire les contraintes suivantes: $x_{21} = x_{22}$ et de $x_{23} = x_{24}$. Il s'agit là d'un élément clé : ***les scénarios avec un historique commun (jusqu'à au moins une étape donnée) devront avoir le même ensemble de décisions jusqu'à cette étape, et cela devra être pris en compte lors de la formulation d'un problème de décision basé sur une structure d'arborescence.***

Notre programme stochastique avec recours (dans le cas de deux périodes) est donc :

Minimiser $2x_1 + \sum_{s=1}^{4} p_s \times (2x_{2s} + 3y_{2s} + 3y_{3s})$

S/c :

$x_1 + y_{2s} >= 500 \qquad s = 1,2$
$x_1 + y_{2s} >= 700 \qquad s = 3,4$

$x_1 + y_{2s} - 500 + x_{2s} + y_{3s} >= 600 \qquad s = 1$
$x_1 + y_{2s} - 500 + x_{2s} + y_{3s} >= 700 \qquad s = 2$
$x_1 + y_{2s} - 700 + x_{2s} + y_{3s} >= 900 \qquad s = 3$
$x_1 + y_{2s} - 700 + x_{2s} + y_{3s} >= 800 \qquad s = 4$

$y_{21} = y_{22}$
$x_{21} = x_{22}$

$y_{23} = y_{24}$
$x_{23} = x_{24}$

Toutes les variables >=0

Récemment, le domaine de la finance quantitative commence à s'inspirer des techniques de programmation stochastique (PS) avec recours pour la résolution de différents problèmes (cf. Zenios [2007]). Pour le cas simple de la décision du choix de la composition d'un portefeuille d'actifs financiers, ces techniques peuvent être utilisées puisque la décision sur la composition doit être prise avant d'observer les performances effectives des classes d'actifs.

Nous considérons également le cas de la gestion de la composition d'un portefeuille d'investissement de façon à obtenir *in fine* une valeur du portefeuille supérieure, ou au moins égale, à un engagement aléatoire (inconnu à la date initiale mais qui sera par la suite connu). Une application des techniques de PS avec recours peut être envisagée : dans le cas mono-périodique par exemple, la décision de départ que prend l'investisseur concerne la composition à t=0 de son portefeuille (l'équivalent de la quantité x_1 à produire dans le cas de la planification de la production). L'évènement aléatoire sera le montant effectif de l'engagement qui ne peut être connu qu'à t=1 (l'équivalent de la demande effective observée à t=1 dans le cas de la planification de la production). Le recours peut être mis en place à travers la liquidation d'une partie du portefeuille (cf. Rasmussen [2004]).

L'application des techniques de PS peut être aussi envisageable dans le domaine de l'assurance. Nous considérons par exemple le cas d'une compagnie d'assurance qui, à la date t=0, encaisse des primes qu'elle prévoit d'investir dans différentes classes d'actifs. Les rendements de cet investissement peuvent être stochastiques ou déterministes (en fonction de la nature des classes d'actifs retenues). D'un autre côté, nous supposons que les prestations à fournir sont stochastiques (cas par exemple de l'assurance bâtiment où la prestation peut augmenter suite à des tempêtes ou à des inondations). Le recours correspond dans ce cas à la possibilité de liquidation (vente) d'une partie des actifs afin de pouvoir faire face à une éventuelle prestation plus élevée que prévu. La question qui se pose est comment la compagnie structure ses décisions d'investissement de manière à utiliser judicieusement ses actifs et de pouvoir honorer sans difficultés ses engagements.

Nous proposons dans ce travail d'appliquer les techniques de PS avec recours pour effectuer l'allocation stratégique d'actifs d'un régime de retraite partiellement provisionné. Cela sera développé de façon détaillée dans un cadre de gestion actif-passif (cf. la section III de ce chapitre). Dans ce qui suit, une formulation mathématique de la PS avec recours sera présentée.

II-4 Formulation mathématique de la programmation stochastique avec recours

Nous partons dans un premier temps de l'étude des programmes stochastiques linéaires (cas de l'exemple développé dans la sous-section précédente) pour ensuite généraliser la démarche. Selon Birge et Louveaux [1997] la forme générale d'un programme linéaire déterministe est la suivante :

$$\begin{aligned} &\min_{x} c^{'} x \\ &s.c.\ Ax \geq b \\ &\quad x \geq 0 \end{aligned}$$

Avec x un vecteur de décision $(n \times 1)$ et c, A et b sont des données connus de taille $(n \times 1)$, $(m \times n)$ et $(m \times 1)$.

Cette formulation suppose que nous avons des informations précises et certaines sur les paramètres de la matrice A et des vecteurs c et b. L'incertitude existante dans le monde de la finance limite l'utilité de cette formulation et conduit à son insuffisance lors de la résolution des problèmes d'optimisation en général.

Le passage à un programme linéaire stochastique tient compte de cette limite et suppose que nous avons de l'aléa au niveau de certaines données ou de certains paramètres. La formulation du problème sera dans ce cas comme suit :

$$\begin{aligned} &\min_x c(\omega)' x \\ &s.c.\ A(\omega)x \geq b(\omega) \\ &\quad x \geq 0 \ \ \omega \in \Omega \end{aligned}$$

Il s'agit d'un programme linéaire paramétrisé par un vecteur d'évènements aléatoires ω (*a priori* continu). Trouver une solution x tel que les contraintes sont toujours satisfaites apparaît être impossible dans ce cas (x est une variable aléatoire fonction de ω) : c'est pour cette raison que les contraintes du programmes ne peuvent pas être présentées sous forme d'égalités. Une approche possible est de « relaxer » les contraintes en acceptant que dans certains cas elles ne seront pas respectées, nous cherchons dans ce cas à limiter la probabilité de ces dépassements. Ceci conduit aux modèles dits « *chance-constrained models* » tel que :

$$\begin{aligned} &\min_x c' x \\ &s.c.\ Ax \geq b \\ &\quad P\{G(\omega)x \geq h(\omega)\} \geq \alpha \\ &\quad x \geq 0 \end{aligned}$$

Nous séparerons par ce modèle les contraintes déterministes des contraintes avec incertitude. La facilité de résolution de ce type de problème dépend de la distribution des paramètres incertains : convexité ou non convexité du modèle (le cas de la convexité rend plus facile la résolution). A part ce problème de calcul, une autre difficulté est potentiellement rencontrée avec ce type de modèle : même s'il permet la prise en compte de l'incertitude existante dans le processus de prise décision, ce modèle n'est pas apte à modéliser des processus de décision dynamiques (où les décisions sont réajustées en fonction des nouvelles informations acquises avec le passage du temps).

Dans un processus de décision dynamique, nous commençons par prendre un ensemble de décisions à la date initiale, basées sur peu d'informations disponibles, mais nous ajustons ensuite ces décisions lorsque l'incertitude a disparue. Cet ajustement conduit bien évidemment à des coûts additionnels, et l'objectif sera de prendre des décisions minimisant aussi bien les coûts immédiats que les coûts futurs espérés des éventuels ajustements. Cette approche nous conduit au principe des modèles de programmation stochastique avec recours.
Le concept de recours peut être défini comme la possibilité de prendre des actions correctives après la réalisation de l'évènement aléatoire. Le cas le plus simple de ce type de modèle est celui avec deux étapes (présenté dans la sous-section précédente) :

- dans la première étape, nous nous disposons que de peu d'information et l'évènement aléatoire n'est pas encore survenu, donc nous décidons la valeur d'une variable de contrôle x.

- dans la deuxième étape, nous nous disposons de l'information certaine sur l'évènement qui était aléatoire, et nous prenons une décision de recours y afin de corriger l'erreur sur la prévision de cet évènement.

Le problème d'optimisation sera formulé de façon à choisir les valeurs optimales de x et y. Notons qu'une propriété importante caractérise ce type de modèles : la décision prise à la première étape (x) est indépendante du scénario pouvant avoir lieu à la deuxième étape : Il s'agit de la propriété de non-anticipation (*nonanticipativity property*). La programmation stochastique avec recours et à deux étapes cherche en général à minimiser le coût de la décision à la première période ainsi que le coût espéré du recours effectué à la deuxième étape.

Nous supposons dans ce qui suit que ω est une variable aléatoire discrète pouvant avoir s valeurs possibles : $\omega_1, \ldots, \omega_s$ avec les probabilités $p_1,.., p_s$ (chaque valeur de ω_i, $i=1,.., s$ correspond à un scénario).

La formulation d'un programme stochastique linéaire avec recours (à deux étapes) peut être présentée comme suit (cf. Brandimarte [2006]) :

$$\begin{aligned} &\min_x \; c'x + E_\omega\left(Q(x,\omega)\right) \\ &s.c. \; Ax = b \\ &\quad x \geq 0 \end{aligned}$$

avec :

$$\begin{aligned} &Q(x,\omega) = \min_y \left(q(\omega)' y\right) \\ &s.c. \; T(\omega)x + W(\omega)y(\omega) = h(\omega) \\ &\quad y \in Y \subset \Re^{n_2} \end{aligned}$$

La fonction de recours $E_\omega\left(Q(x,\omega)\right)$ est fournie par les quatre éléments suivants :

- un ensemble Y qui décrit l'ensemble réalisable des actions de recours, par exemple $Y = \left\{y \in \Re^{n_2} \;/\; y \geq 0\right\}$;
- q : un vecteur de coûts de recours ($n_2 \times 1$);
- W : une matrice $s \times n_2$, elle est appelée la matrice de recours.
- T : une matrice $n_2 \times n$ reliant x à y, elle est appelée la matrice de technologie (nous rappelons que x est un vecteur de décision de taille $n \times 1$).

La formulation présentée ci-dessus est composée de deux programmes. Le premier programme minimise les coûts directs de la première étape, $c'x$, plus l'espérance du coût du recours appelée également fonction de recours, $E_\omega\left(Q(x,\omega)\right)$. Cela tient compte évidemment de toutes les réalisations possibles de l'évènement aléatoire ω et du respect de la contrainte $Ax = b$. Le coût de recours $Q(x,\omega)$ dépend à la fois de la valeur x (la décision de la première étape) et de l'élément aléatoire ω. Le deuxième programme décrit comment choisir les décisions qui forment $y(\omega)$. L'objectif à ce stade est de minimiser le coût de recours $q(\omega)' y$

compte tenu d'une contrainte liée aux actions de recours elles-mêmes : $T(\omega)x+W(\omega)y(\omega)=h(\omega)$.

Le cas général suppose que $Q(x,\omega)=\min_{y}(q(y,\omega))$. Il est souvent considéré que $q(y,\omega)$ est une fonction linéaire ayant pour forme $q(\omega)' y$ et que y est non-négatif (cf. Castro [2009]). Le cas de la non-linéarité de $q(y,\omega)$ peut être envisagé. De telles fonctions sont difficiles à traiter puisque cela implique l'intégration multidimensionnelle d'une fonction définie implicitement à travers un problème d'optimisation. Même dans le cas où nous ne disposons pas d'une forme analytique simple de $q(y,\omega)$, le fait que dans la plupart des cas la fonction de recours est convexe rend possible l'évaluation (ou l'estimation) de ses valeurs et de trouver un sous-gradient à un point donné x (cf. Brandimarte [2006]).

Nous pouvons réécrire la formulation ci-dessus sous un seul et unique programme, appelé la forme étendue (*extensive form*) :

$$\begin{aligned}
&\min_{x}\ c'x + E_{\omega}(Q(x,\omega)) \\
&s.c.\ Ax=b \\
&\quad T(\omega)x+W(\omega)y(\omega)=h(\omega) \\
&\quad x\geq 0, y\geq 0
\end{aligned}$$

En supposant que pour chaque scénario il existe une action de recours correspondante (y_i , i =1,.., s), le problème de programmation stochastique linéaire avec recours (à deux étapes) se présentera comme suit :

$$\begin{aligned}
&\min_{x,y_i}\ c'x+\sum_{i=1}^{s} p_i q'(\omega_i) y_i \\
&s.c. \\
&\left.\begin{aligned}
&Ax=b \\
&T(\omega_i)x+W(\omega_i)y_i=h(\omega_i) \\
&y_i\geq 0 \\
&x\geq 0
\end{aligned}\right\} i=1,\ldots,s
\end{aligned}$$

Les éléments développés jusque là restent valables dans le cas d'un programme stochastique non-linéaire (à deux étapes). Le problème peut être présenté sous la forme générale suivante :

$$\begin{aligned}
&\min_{x,y_i}\ c(x)+\sum_{i=1}^{s} p_i q(y_i;\omega_i) \\
&s.c. \\
&\left.\begin{aligned}
&T(\omega_i)x+W(\omega_i)y_i=h(\omega_i) \\
&y_i\in Y \\
&x\in X\subset \Re^n
\end{aligned}\right\} i=1,\ldots,s
\end{aligned}$$

où $c(x)$: correspond au coût engendré par les décisions de la première étape.

X : reflète l'ensemble des contraintes à respecter par x (dans le cas linéaire étudié

précédemment, X a été définie comme suit : $X = \{x : Ax = b, x \geq 0\}$)

Finalement, nous pouvons aller plus loin pour généraliser cette présentation au cas multi-périodique (ou multi-étapes). Dans ce cas, nous nous trouvons face à une séquence temporelle du couple observation/décision. Pour k étapes, il est possible de formuler le problème ainsi (cf. Castro [2009]) :

$$\min_{x_i} \; c_1(x_1) + E_{\omega^2}\left[\min c_2(x_2(\omega^2);\omega^2) + ... + E_{\omega^k}\left[\min c_k(x_k(\omega^k);\omega^k)\right]..\right]$$

$$\begin{aligned} s.c. \quad & W_1 x_1 = h_1 \\ & W_2 x_2(\omega^2) + T_1(\omega^2) x_1 = h_2(\omega^2) \\ & W_i x_i(\omega^i) + T_{i-1}(\omega^i) x_{i-1}(\omega^{i-1}) = h_i(\omega^i) \qquad i = 3,...,k \\ & x_1 \geq 0, x_i(\omega^i) \geq 0 \qquad i = 2,...,k \end{aligned}$$

Avec :

W_i ($i = 1...,k$) : représentent des matrices connues de taille $m_i \times n_i$.

h_1 : représente un vecteur connu dans $\Re^{m_1}$.

h_i ($i = 2...,m_i$) : représentent des vecteurs aléatoires dans $\Re^{m_i}$.

T_i ($i = 1...,k-1$) : représentent les matrices aléatoires de taille $m_i \times (n_i - 1)$.

x_i ($i = 1...,k$) : représente le vecteur des décisions à l'étape i avec $x_i \in \Re^{n_i}$ (pour $i > 1$, x_i dépend des évènements aléatoires antérieures).

ω^i : représente l'historique des évènements aléatoires jusqu'à l'étape i.

La formulation ci-dessus suppose implicitement que les décisions prises dépendent exclusivement des données du passé : ces décisions sont les mêmes à l'étape i pour les scénarios qui ont un historique commun jusqu'à l'étape $i-1$. Pour une formulation déterministe équivalente, ces contraintes « non-antipatives » doivent être prises en compte (cf. Castro [2009]).

Après avoir étudié le cadre général de modélisation mathématique des PS avec recours, nous nous intéressons dans la sous-section suivante à la panoplie d'outils disponibles pour la résolution de tels problèmes.

II-5 Résolution des programmes stochastiques

Pour certains problèmes complexes de programmation stochastique, les méthodes classiques de résolution (tel que le simplexe et la méthode de point intérieur) s'avèrent parfois inefficaces pour trouver une solution exacte ou dans un temps de calcul raisonnable. Nous devons alors recourir à des algorithmes d'optimisation dédiés. Plusieurs méthodes peuvent être envisagées suivant la nature et la complexité des problèmes rencontrés (dont certaines ont été présentées dans le cadre de l'étude des techniques d'optimisation numérique, cf. section II du chapitre 2 de la partie II).

La méthode la plus répandue est l'algorithme de séparation et d'évaluation (« *Branch and Bound* », cf. Lawler et al. [1966]). Il s'agit d'une méthode d'énumération implicite des solutions. Lors de l'étape de séparation (« *branching* »), les solutions du problème sont énumérées. Cette étape permet de diviser la région réalisable en sous-régions, pouvant à leur

tour être divisées récursivement en d'autres sous-régions. Chaque sous-région possède un ensemble de solutions réalisables. Un arbre de solutions est ainsi construit, formant l'arbre de décision ou l'arbre de recherche.

L'étape d'évaluation (« *bounding* ») intervient pour éviter l'énumération systématique de toutes les solutions. Les solutions trouvées sont utilisées dans la phase d'évaluation pour définir des bornes supérieures et inférieures de la solution optimale. L'algorithme s'arrête généralement suivant une contrainte imposée sur le temps de calcul ou sur la précision de la solution. La plupart des outils commerciaux d'optimisation reposent sur cette méthode de résolution.

D'autres méthodes s'appuient sur l'algorithme de séparation et d'évaluation, mais permettent de restreindre astucieusement l'espace des solutions par l'introduction de contraintes supplémentaires au problème initial. Elles réfèrent généralement aux méthodes de décomposition (« *branch and cut* »). La plus connue d'entre elles est la décomposition de Benders (Dogan et Goetschalckx [1999]). Cette méthode repose sur la division d'un problème en un problème maître et en un ensemble de sous-problèmes. Des solutions sont successivement obtenues du problème maître et corrigées, au besoin, par les variables duales calculées des sous-problèmes pour ainsi ajouter une coupe d'optimalité au problème maître. L'algorithme s'arrête sur la base d'un critère d'arrêt sur les temps de calcul ou sur la précision de la solution. Dans un contexte stochastique, son équivalent est désigné par la méthode de décomposition « *L-Shaped* » (*L-Shaped decomposition method*, cf. Santosa et Alexandrov [2005]).

Certains problèmes peuvent être résolus efficacement par la relaxation lagrangienne (Pirkul et Jayaraman [1996]). Cette méthode consiste à modifier le problème initial en introduisant certaines des contraintes à la fonction objectif. Ces contraintes, pondérées d'un multiplicateur de Lagrange, ajoutent une pénalité à la fonction objectif si elles ne sont pas satisfaites. La difficulté de cette méthode de résolution consiste à déterminer la valeur des multiplicateurs.

Les méta-heuristiques sont abordées en dernier recours lorsque ces dernières méthodes de résolution se sont avérées vaines. Les méta-heuristiques[20] sont des algorithmes itératifs qui progressent vers un optimum par échantillonnage d'une fonction objectif. Dans un espace de recherche, une ou plusieurs solutions sont évaluées et comparées afin d'éliminer les solutions de mauvaise qualité et ainsi arriver vers une solution optimale approximée. Une borne inférieure et supérieure de la solution est calculée dans cette démarche. L'algorithme prend fin lorsqu'un critère d'arrêt est rencontré, soit en fonction du temps de calcul ou encore de la précision obtenue des bornes calculées. Pour rappel, certains exemples de ces algorithmes ont étaient étudiés dans le cadre des techniques d'optimisation numérique pour l'allocation d'actifs (cf. section II du chapitre 2 de la partie II).

Parmi les méta-heuristiques les plus connues, notons les algorithmes évolutionnaires incluant les algorithmes génétiques (Alliot [1996]), le recuit simulé (Branke et al. [2008]) et la recherche avec tabous (Mazzola et Schantz [1997]). Une méthode a été proposée plus spécifiquement à l'égard de modèles stochastiques avec recours faisant intervenir des facteurs aléatoires indépendants. Il s'agit de la méthode d'approximation de la moyenne d'échantillonnages (*Sample Average Approximation* – SAA), basée sur les techniques d'échantillonnage de *Monte Carlo* (Santosa et al. [2005]).

[20] Pour rappel, certains exemples de ces algorithmes ont été étudiés dans le cadre des techniques d'optimisation numérique pour l'allacation d'actifs (cf. section V, chapitre 2 de la partie II)

Cette méthode permet de contourner les difficultés d'évaluation des probabilités d'occurrence des scénarios. Des échantillons de scénarios sont définis hors de la procédure d'optimisation, suivant les fonctions de distribution considérées pour les facteurs aléatoires. Une ou des solutions sont d'abord obtenues d'échantillons faisant intervenir un nombre plus restreint de scénarios. La réaction des solutions face aux facteurs aléatoires est ensuite étudiée avec un échantillon de plus grande taille.

Pour les deux étapes, la partie des coûts reliée aux décisions sujettes aux facteurs aléatoires se calcule par la moyenne des coûts pour chaque scénario. La méthode d'approximation de la moyenne d'échantillonnages incorpore une procédure d'analyse de convergence. Une borne inférieure et borne supérieure est ainsi calculée à l'égard de la solution optimale du problème (Shapiro et al. [2003]). Ces bornes servent à sélectionner la meilleure solution et évaluer la qualité du résultat.

La plupart des méthodes exposées ci-dessus, allant de l'algorithme de séparation et évaluation jusqu'à la méthode d'approximation de la moyenne d'échantillonnages, sont implémentées sous forme de solveurs et utilisées dans des outils plus globaux d'optimisation numérique (cf. Brandimarte [2006] et Klassen [1997]). Cela concerne également le solveur que nous allons utiliser dans la section suivante, dédiée à l'exploitation des techniques de PS dans le cadre de l'allocation stratégique d'actifs d'un régime de retraite typique.

III- Nouvelle approche d'ALM par la discrétisation des scénarios économiques

Il s'agit ici de proposer notre conception pour l'adaptation des techniques de programmation stochastique, en particulier celles avec recours, dans un contexte d'allocation stratégique d'actifs et d'ALM d'un régime de retraite partiellement provisionné. Nous décrivons, au cours de la même proposition, une nouvelle méthodologie de génération de scénarios économiques que nous appelons méthodologie « des quantiles de référence ».

Tout au long de l'application numérique développée ci-après, nous comparons des résultats relatifs aux deux stratégies d'allocation d'actifs : celle basée sur la stratégie *Fixed-Mix* et celle basée sur les techniques de programmation stochastique. Nous testons également la sensibilité de l'approche développée d'ALM par rapport au changement de certains de ses paramètres, toutes choses étant égales par ailleurs.

III-1 Méthode des quantiles de référence pour le GSE

L'utilisation des techniques de programmation stochastique nécessite de représenter les incertitudes futures en utilisant une structure arborescente (*event tree*) qui illustre de façon intuitive l'aléa (cf. section II, chapitre 1 de la partie I). Au-delà de cette contrainte d'ordre technique, la structure arborescente a le mérite de synthétiser le problème de projection des scénarios économiques et financiers et de simplifier son interprétation par le décideur : il s'agit d'une représentation plus simplifiée du problème initial où la projection est effectuée avec une structure linéaire (cf. section II, chapitre 1 de la partie I). Plus précisément, nous réduisons la dimension du problème en passant d'une structure linéaire sous forme de matrice $(n \times m)$[21] à une structure arborescente avec un nombre plus petit de nœuds. Un autre atout

[21] n pour le nombre de simulation et m pour le nombre de pas (fonction de l'horizon de projection et du pas choisie)

consiste dans l'aspect conditionnel qui caractérise le lien entre les nœuds-enfants de la structure arborescente à une date donnée et les nœuds de la date précédente : cela permet la prise en compte, dans le temps, des relations entre les différentes valeurs projetées.

Pour rappel, un exemple de cette structure est présenté dans le graphique 73. Dans le graphique 73, la structure de l'arbre comporte 64 scenarios. Nous retenons 8 réalisations possibles durant la première période. A partir de chacune de ces réalisations, nous déterminons 4 réalisations durant la deuxième période. Chacune d'entre elle est ainsi appelée « réalisation conditionnelle » car elle est fonction de la valeur dans le nœud antécédent. La troisième période est caractérisée par la génération de 2 réalisations à partir de celles obtenues à la fin de la deuxième période. Ceci donne un total de $8 \times 4 \times 2 = 64$ scénarios (ou trajectoires).

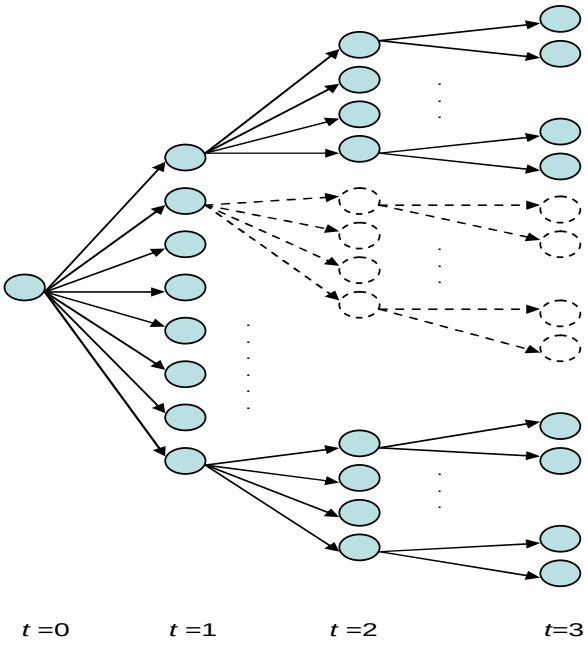


Fig. 73 : *Exemple de la structure schématique d'arborescence pour les trajectoires simulées*

Dans cette étude sur l'application des techniques de la programmation stochastique, le générateur d'arbre se base sur une méthodologie que nous appellerons méthodologie « des quantiles de référence ». Cette dernière repose sur deux éléments clés : le choix d'une variable, parmi celles modélisées, comme variable de référence (ex. les rendements des actions) et le passage de la structure linéaire de projection à une structure arborescente (pour chacune des variables modélisées) en se basant sur un certain nombre de quantiles de la variable de référence. La méthodologie « des quantiles de référence » passe par les étapes suivantes :

- Dans un premier temps, nous générons un nombre élevé de scénarios pour chacune des variables modélisées selon le principe de la structure linéaire et compte tenu de la période de projection (ex. 1 000 trajectoires possibles sur les 20 prochaines années avec un pas annuel, cf. sous section II-1, chapitre 1 de la partie I)
- ensuite, nous sélectionnons une classe d'actifs ou une variable donnée (ex. les rendements des actions) qui sera appelée « variable de référence ».
- puis, à chaque niveau de l'arbre (date t =1 par exemple) associée à la variable

de référence (rendement des actions), nous retiendrons un nombre prédéfini de quantiles « quantiles de référence » en fonction de la structure de l'arbre, en particulier en fonction du nombre des nœuds prévu à cette date (ex. les trois quantiles 99,5 %, 50 % et 0,5 % des rendements des actions à la date *t* =1 si le nombre de nœud à cette date est prévu être égal à 3).

Par ailleurs, nous notons que si le nombre d'années considérées entre deux niveaux de l'arbre est supérieur à 1, le générateur de l'arbre déduit le rendement simulé sur la période considérée (ex. 5 ans) à partir du générateur linéaire mentionné ci-dessus selon la formule suivante :

$$R_{arbre} = \prod_{i=1}^{H} (1 + R_i)$$

où :

H est la période considérée dans l'arbre (ex. 5 ans)

R_{arbre} est le rendement sur la période H

R_i est le rendement annuel simulé dans la structure linéaire

- Les projections sous forme d'arbres des autres variables (en dehors de la variable de référence) sont ensuite déterminées : en chaque nœud de ces arbres, nous donnons à ces variables la valeur qui correspond à la variable de référence, sur la base des scénarios simulés en amont avec la structure linéaire de projection. Cette démarche a le mérite de conserver la structure de dépendance des scénarios générés initialement. Nous obtenons ainsi, pour chacune des variables modélisées (les rendements des obligations, l'inflation, etc.) son propre arbre de scénarios.
- Les flux de passif de départ, exprimés en euros désinflatés, seront également inflatés, compte tenu de l'arbre obtenu pour l'inflation : ainsi ces flux inflatés seront également représentés sous forme d'arbre.

La question du lien entre le choix d'une part de n quantiles (q_i) dans l'arbre des scénarios (ce qui revient au choix des niveaux de probabilité cumulée α_i qui leur sont associés) et d'autre part le choix des poids (p_i) correspondant dans l'arbre des scénarios se pose au niveau de l'étape 3 du schéma présenté ci-dessus. Le poids (p_i) de chaque quantile reflète la probabilité de son observation par rapport à celles du reste des quantiles observés à la même date. Il s'agit du poids donné au quantile q_i dans l'arbre en fonction de la probabilité de son observation. Nous nous référons au principe d'intégrale de Riemann afin de pouvoir mettre en évidence ce lien.

Considérons le cas d'une variable aléatoire X avec une fonction de répartition F. Nous supposons aussi que la fonction inverse F^{-1} existe et est continue. L'objectif est de construire la variable discrète $\hat{X}_n$ à travers le choix de n quantiles $(q_1, q_2, ..., q_n)$ tel que $q_i = F^{-1}(\alpha_i)$ pour $i = 1,...,n$ et α_i représentant le niveau de probabilité cumulée associé au quantile q_i. Cette variable $\hat{X}_n$ doit également refléter la distribution initiale de X dans le sens qu'au moins les espérances des deux variables doivent converger : $E(\hat{X}_n) \underset{n \to \infty}{\approx} E(X)$.

Il s'agit d'une hypothèse simpliste de correspondance entre les deux distributions (celle de X et celle de $\hat{X}_n$) mais qui permet à ce niveau de formuler une première étude sur le lien entre d'une part le poids p_i, associé à chaque quantile, et d'autre part le niveau de probabilité α_i. Nous cherchons donc les quantiles (q_i) et les poids (p_i) associés à ces quantiles de sorte que la somme pondérée des n quantiles convergent vers $E(X)$ lorsque n tend vers l'infini.

Pour cela, nous supposons que U représente la loi uniforme sur l'intervalle [0,1], nous pouvons écrire donc : $E(X) = \int_0^1 F^{-1}(u)du$. D'un autre côté, l'espérance de la variable $\hat{X}_n$ se présente comme : $E\left(\hat{X}_n\right) = \sum_{i=1}^{n} p_i q_i = \sum_{i=1}^{n} p_i F^{-1}(\alpha_i)$. Le calcul de cette dernière revient donc à un calcul de l'intégrale de Riemann pour la fonction F^{-1}. Autrement dit, la valeur de $E\left(\hat{X}_n\right)$ peut être obtenue à travers la somme des rectangles de base égale à la longueur du segment dont le milieu est α_i (segment que nous pouvons appelé I_i) et de hauteur $F^{-1}(\alpha_i)$. Le graphique 74 illustre cette approche. Le poids p_i peut être estimé à travers la longueur de la base du rectangle contenant α_i (ou la longueur de I_i).

Dans un premier temps, nous pouvons remarquer à travers le graphique 74 que p_i est fonction de la position et donc du niveau des α_i ($i = 1,...,n$). Ainsi, le fait de supposer un α_i ($i = 1,...,n$) régulier, c'est à dire avec une distance identique entre les différentes valeurs de α_i, permet d'obtenir *in fine* un poids $p_i = 1/n$ (cf. graphique 75). En conséquence, avec une valeur élevée de n, l'hypothèse de régularité de α_i garantit la convergence des deux espérances : $E(\hat{X}_n) = \sum_{i=1}^{n} p_i q_i = \sum_{i=1}^{n} \frac{1}{n} q_i \underset{n\to\infty}{\approx} E(X)$. Le choix de n valeurs de α_i équidistants nous autorise à estimer l'espérance par une somme pondérée où chaque p_i est égale à $1/n$. Autrement dit, pour une date donnée, si nous optons pour des quantiles équidistants, le choix de poids identiques pour chaque quantile permet de refléter la distribution initiale de X.

Considérons maintenant le cas de plusieurs variables que l'on veut représenter sous forme d'arbre de scénarios tout en conservant les liens entre elles : retenons en particulier l'exemple de trois variables X, Y et Z avec X comme variable de référence. L'objectif est de déterminer le triplet d'observations (x_i, y_i, z_i) à un instant donné et à un nœud précis de l'arbre des scénarios à partir des projections effectuées selon la structure linéaire. Une approche possible serait de commencer par la discrétisation de la variable de référence X comme indiqué ci-dessus pour ensuite retenir les observations simultanément simulées de Y et Z (sur la base d'un modèle de corrélation prédéfini dans la structure linéaire de projection). Cela se rapproche du schéma décrit dans la méthodologie des quantiles de référence : après avoir déterminé le rang de x_i (ex. 50ème simulation dans la structure linéaire), nous retenons les valeurs de y_i et de z_i correspondant à la même simulation i.

Outre la prise en compte de la structure de dépendance entre les variables, la méthodologie « des quantiles de référence » permet de lier les atouts d'une structure linéaire (cf. sous section

II-1, chapitre 1 de la partie I) à ceux de l'utilisation d'une structure arborescente. La méthodologie proposée réduit également la dimension du problème initial : le nombre de scénarios dans l'arbre obtenu avec cette méthodologie peut être beaucoup moins élevé que celui dans le cas de la structure linéaire (situé, souvent, au-delà de 1 000 scénarios pour avoir une bonne estimation de la distribution des variables considérées).

Une autre alternative est également envisageable : il s'agit de raisonner en termes d'espérance d'un intervalle J_i des points limités par deux quantiles successifs (cf. graphique 74) au lieu de raisonner en termes de quantiles. Dans ce cadre, nous aurons :

$$\hat{Y}_n = \hat{Y}(\alpha_n) = E[Y / X \in J_n] \qquad (i = 1,...,n)$$

$$\hat{Z}_n = \hat{Z}(\alpha_n) = E[Z / X \in J_n] \qquad (i = 1,...,n)$$

Cela permet d'avoir *in fine* une estimation, basée sur d'avantages d'observations, du quantile représentant la distribution initiale.

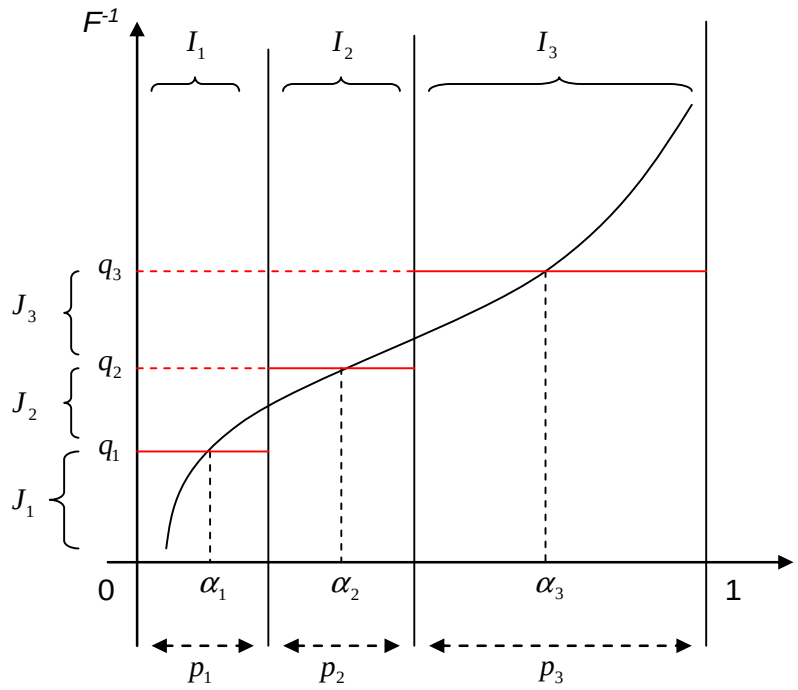


Fig.74 : *Illustration du lien entre le niveau de probabilité cumulée α_i d'un quantile q_i et le poids associé à ce dernier dans l'arbre des scénarios (cas où n =3)*

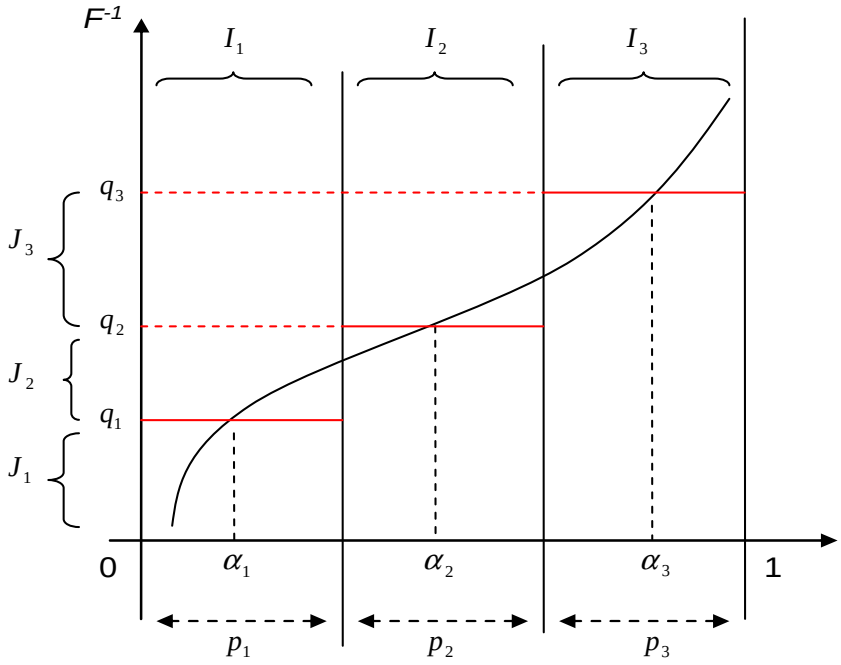


Fig.75 : *Illustration du lien entre le niveau de probabilité cumulée* α_i *d'un quantile* q_i *et le poids associé à ce dernier dans l'arbre des scénarios (cas où* $n=3$ *et* α_i *est régulier) :* $p_1 = p_2 = p_3$

Ces approches s'insèrent dans le cadre de la recherche d'une vision à la fois simplifiée, réelle et dynamique des stratégies possibles pour l'allocation stratégique d'actifs, en présence de contraintes d'ALM. La méthodologie des quantiles de référence tient compte des hypothèses de distribution des différentes variables projetées : elle garde la structure de risque initiale en retenant les quantiles « approximatifs » de cette distribution. Il s'agit d'une stratégie à la fois utile et cohérente. Sa mise en œuvre n'est pas compliquée et elle peut être adaptée pour étudier d'autres problématiques.

III-2 Modèle d'optimisation basé sur la programmation stochastique

Par analogie au principe de la programmation stochastique avec recours (développé dans les sous-sections précédentes), nous pouvons considérer l'action de recours dans un modèle d'ALM d'un régime de retraite provisionné comme la possibilité de liquider une partie du portefeuille de marché compte tenu de la réalisation de l'évènement aléatoire (ω) qui affectera dans notre cas aussi bien l'actif que le passif du régime, spécificité issue du contexte d'ALM (ex. performance des actions côté actif, inflation constatée des salaires côté passif). La décision de début de période concerne donc la répartition des poids des différents actifs du portefeuille de marché (x). Elle va être ajustée à la fin de la même période en fonction des réalisations de l'évènement aléatoire considéré. Le cadre d'une allocation dynamique d'actifs est ainsi vérifié.

Pour le modèle d'optimisation, nous présentons dans un premier temps un cadre général et typique de problème d'optimisation que différents types de régimes de retraite peuvent rencontrer, notamment ceux partiellement provisionnés (cf. section IV de la partie II). Ce modèle sera par la suite simplifié au niveau de l'implémentation. Le but est ici de proposer une approche générale permettant la mise en place d'un modèle d'ALM dynamique basé sur la programmation stochastique multi-périodiques.

La programmation stochastique est caractérisée par l'utilisation d'une fonction objectif qui devra être optimisée et d'un ensemble de contraintes exprimées sous forme d'équations linéaires ou non linéaires et qui doivent être respectées. La fonction objectif est introduite afin qu'elle illustre la fonction d'utilité du décideur. La satisfaction des contraintes pourrait être obtenue par une multitude de solutions. La PS cherche à déterminer la meilleure solution parmi celles qui satisfont les contraintes. C'est le cas d'une analyse type ALM où la satisfaction de l'adossement des flux de passifs par les flux d'actifs peut être faite avec différentes allocations de portefeuille. La formulation de ce problème à travers la PS, permet d'aboutir aux décisions d'allocations d'actifs optimales dans le cadre de l'ALM.

Le fonctionnement financier du régime (ou du fonds) de retraite en France, objet de cette étude, peut être présenté comme suit : les flux de cotisations permettent de régler les flux de prestations. Le surplus, le cas échéant, permet d'alimenter une réserve destinée à régler une partie des prestations futures. Cette même réserve peut se voir prélever, le cas échéant, le solde technique débiteur (cotisations insuffisantes pour régler les prestations). Entre temps, la réserve est placée sur le marché financier et répartie entre différentes classes d'actifs.

De façon générale, nous considérons le cas d'un fonds de retraite qui possède à la date initiale t_0 une réserve W_0 et qui devra faire face à une série de flux F_t (t =1,…. T) sur les T prochaines années. F_t est positive dans le cas où les cotisations sont supérieures aux prestations et elle est négative dans le cas contraire. Nous supposerons que le dernier flux F_T est négatif et qu'il représente l'engagement final du fonds vis-à-vis de ses adhérents.

Nous nous intéressons à la stratégie d'investissement qui permet de parvenir au moins à payer ces différents flux, en particulier le dernier F_T (ou L). L'idéal serait, bien sûr, d'avoir un montant de réserve finale supérieur à F_T. Ceci étant dit, la prise en compte de l'aversion au risque est elle aussi nécessaire puisque nous pouvons se retrouver avec une réserve finale W_T inférieure à F_T.

La fonction objectif considérée est une fonction d'utilité linéaire : le fonds de retraite attribue un malus q pour chaque unité de déficit (écart négatif entre le montant de la réserve finale W_T^s et l'engagement final F_T^s pour le scénario s) et un bonus r pour chaque unité de surplus. Pour un scénario s avec un écart positif entre W_T^s et F_T^s, l'utilité totale du fonds de retraite sera augmentée de $r.\left(W_T^s - F_T^s\right).p^s$ (avec p^s la probabilité associée au scénario s). Pour un scénario s avec un écart négatif entre W_T^s et F_T^s, l'utilité totale du fonds de retraite sera diminuée de $q.\left(W_T^s - F_T^s\right).p^s$. L'utilité totale du fonds sera supposée être égale à la somme de ces quantités calculées pour tous les scénarios projetés.

Côté contraintes, différents aspects peuvent être considérés. La contrainte principale sera appelée la contrainte de probabilité de ruine : il s'agit d'assurer que la stratégie optimale d'allocation d'actifs permet d'avoir un W_T^s supérieur à F_T^s dans (1- p_{ctrt}) fois du nombre total de scénarios. Plusieurs autres contraintes peuvent être considérées à savoir :

- Contraintes budgétaire : le montant investi à chaque instant t dans l'actif j doit être égal au montant investi dans le même actif à t-1 et capitalisé entre ces deux dates augmenté (diminué) du montant des achats (ventes) dudit actif effectués à la date t.

- Contraintes de rebalancement : les flux encaissés à la date t en provenance de la vente de l'actif j ou d'un solde technique positif (cotisations > prestations) constituent l'unique source de financement pour l'achat dudit actif à la même date t (y compris les différents frais de transaction).
- Contraintes de portefeuille : le montant investi dans l'actif j est borné à chaque instant t entre deux montants proportionnels à la valeur de la réserve observée à la même date.
- Contraintes de transaction : les montants des achats et des ventes de l'actif j sont plafonnés à chaque instant t en fonction de la valeur de la réserve observée à la même date.

Le modèle général d'optimisation considéré peut ainsi être présenté comme suit :

Paramètres Déterministes :

h_j^0 : Montant initial investi dans l'actif j

W_0 : Réserve initiale totale

c_j^p : Coût de transaction pour l'achat de l'actif j

c_j^s : Coût de transaction pour la vente de l'actif j

l_j : Poids inférieur du montant investi dans l'actif j (exprimée en pourcentage de la réserve totale).

u_j : Poids supérieur du montant investi dans l'actif j (exprimée en pourcentage de la réserve totale).

b_j^P : Poids maximal du montant d'achat périodique de l'actif j (exprimée en pourcentage de la réserve totale).

b_j^s : Poids maximal du montant de vente périodique de l'actif j (exprimée en pourcentage de la réserve totale).

τ_t : Durée de la période $[t-1, t]$ en nombre d'années

$card(S)$: Nombre total de scénarios (nombre de branches à la date finale T)

p^s : Égale à $1/card(S)$ (hypothèse approximative de la probabilité de chaque scénario)

q : Bonus de chaque unité de surplus (rémunération)

r : Malus de chaque unité de déficit (pénalité)

Paramètres Stochastiques :

$R_{t,j}$: Rendement de l'actif j sur la période [t-1, t],

F_t : Flux de trésorerie de la période [t-1, t]

L_s : Engagement final cible (à la date T) dans le cas du scénario s

Variables de décision :

$h_{t,j}$: Montant investi dans l'actif j entre la date t et $t+1$

$p_{t,j}$: Montant d'achat de l'actif j à la date t

$s_{t,j}$: Montant de vente de l'actif j à la date t

W_t : Réserve totale à la date t

w_+^s : Surplus (écart positif entre la réserve finale et l'engagement final L_s pour le scénario s)

w_-^s : Déficit (écart négatif entre la réserve finale et l'engagement final L_s pour le scénario s)

$card(W_-)$: Nombre total des scénarios où l'écart entre la réserve finale et l'engagement final est négatif (nombre de scénario où w_-^s >0)

p_{ctrt} : Probabilité maximale acceptée d'avoir des scénarios où la réserve finale est inférieure à l'engagement final (w_-^s >0). Autrement dit, le rapport $card(W_-)/card(S)$ doit être inférieur à p_{ctrt}.

Objectif :

$$Maximiser \sum_{s\in S} p^s\left(qw_+^s - rw_-^s\right)$$

Contraintes :

Contraintes budgétaire

$$\sum_{j\in J} h_{0,j} = W_0$$

$$h_{0,j} = h_j^0 + p_{0,j} - s_{0,j}$$

$$h_{t,j} = R_{t,j}h_{t-1,j} + p_{t,j} - s_{t,j}$$

Contraintes de rebalancement

$$\sum_{j\in J}\left(1+c_j^p\right)p_{0,j} = \sum_{j\in J}\left(1-c_j^s\right)s_{0,j}$$

$$\sum_{j\in J}\left(1+c_j^p\right)p_{t,j} = \sum_{j\in J}\left(1-c_j^s\right)s_{t,j} + F_t$$

Contraintes de portefeuille

$$l_jW_t \le h_{t,j} \le u_jW_t \text{ (Avec : } \sum_j h_{t,j} = W_t)$$

Contraintes de transaction

$$p_{t,j} \le \tau_t b_j^P W_t$$

$$s_{t,j} \le \tau_t b_j^s W_t$$

Contraintes d'équilibre final

$$W_T{}^s = \sum_{j\in J} R_{T,j}h_{T-1,j} = L_s + w_+^s - w_-^s$$

Contraintes de probabilité de ruine

$$card(W_-)/card(S) \le p_{ctrt}$$

$$\forall t = 1,....,T-1, j \in J$$

Pour le modèle d'optimisation, objet de l'implémentation, certains paramètres seront considérés comme nuls afin de simplifier la mise en place, en particulier les paramètres relatifs aux contraintes de transaction, de rebalancement et de portefeuille. Autrement dit, nous supposons l'absence de frais de transaction, des contraintes de rebalancement et de portefeuille définies ci-dessus.

Concernant la mise en place de la méthodologie des « quantiles de référence », nous avons supposé que le nombre minimum de quantiles par niveau (ou date) sera de deux : 97,5 % et 2,5 %. L'ensemble des quantiles supplémentaires sera choisi de façon à ce que les distances, en termes de niveaux de probabilité cumulée, entre les différents quantiles successifs soient les mêmes. Autrement dit, si n_q=5 désigne le nombre de quantiles retenus à une date donnée (y compris celui du 97,5 % et celui du 2,5 %) alors cette distance est égale à :

$$\frac{97{,}5-2{,}5}{n_q-1}=\frac{97{,}5-2{,}5}{5-1}=23{,}75$$

ce qui revient à considérer successivement les quantiles avec les niveaux de confiance suivants : 2,5 %, 26,25 % (= 2,5 %+23,75 %), 50 % (= 26,25 %+23,75 %), 73,75 % (= 50 %+23,75 %), 97,5 %.

Nous considérons aussi qu'à chaque date, les différents quantiles de l'arbre auront la même probabilité d'occurence et ainsi un identique p^s pour tous les scénarios : $p^s = 1/card(S)$. Cela garantit la convergence vers la distribution réelle de la variable modélisée du moment où les niveaux de probabilité cumulée des quantiles sont équidistants (comme montré dans la sous-section précédente).

Dans l'implémentation numérique, la génération des différents arbres est effectuée via le logiciel MATLAB et compte tenu d'un même GSE à structure linéaire identique, en termes de modèles et de paramètres, à celui retenu précédemment dans la sous-section I-2, chapitre 2 de la partie II. L'horizon des projections (et donc des différents flux de passif) est de 20 ans. Le graphique ci-dessous représente la série des flux de passif considérée :

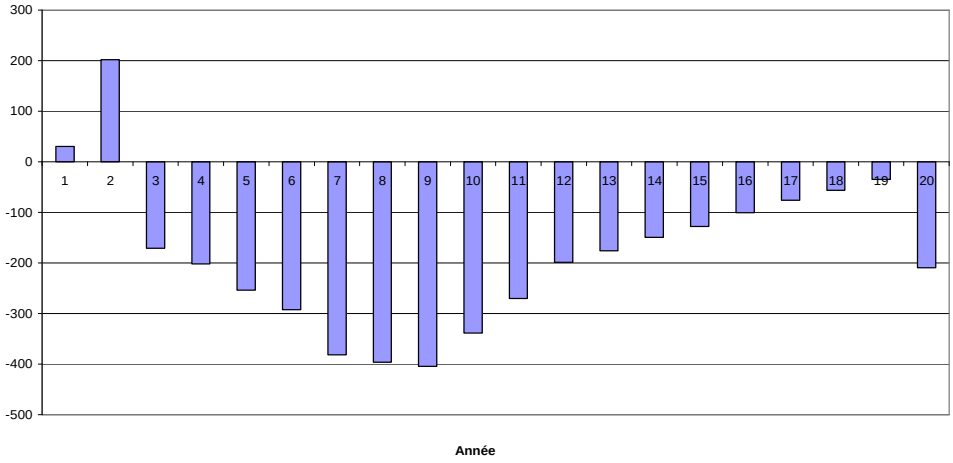


Fig. 76 : *Flux de passif (désinflatés) sur 20 ans*

Parmi les *inputs* supplémentaires de MATLAB nous choisissons, de façon arbitraire, les paramètres standards suivants :

Structure des nœuds de l'arbre	[1 5 4 4]
Structure des périodes de l'arbre	{1 4 15}
p_{ctrt}	20 %
q	2
r	1
W_0	10 000
Variable de référence	Rdt des actions

Tab.39 : *Paramètres standards dans le modèle d'optimisation stochastique multi- périodique*

Nous rappelons que la structure des nœuds de l'arbre correspond au vecteur du nombre de nœuds-enfants à chacune des quatre niveaux considérés dans l'exemple ($[n_{t_0} n_{t_1} n_{t_2} n_{t_3}]$). La structure des périodes de l'arbre correspond au vecteur des horizons de projection supposés entre deux niveaux ($\{H_1\ H_2\ H_3\}$).

L'implémentation du modèle d'optimisation stochastique multi-périodique est quant à elle effectuée via le langage de modélisation AMPL (cf. Fourer et al. [2002]). Ce dernier est adapté pour le type de problématique basé sur la programmation stochastique avec recours : après lui avoir précisé les différents paramètres (les paramètres déterministes, les paramètres stochastiques, les variables de décisions, l'arbre des scénarios, etc.), l'optimisation est effectuée avec un solveur dédié à la PS avec recours (cf. Bandimarte [2006]).

Parallèlement, le problème d'optimisation sera mis en place dans le cadre de la stratégie *Fixed-Mix* en gardant le même GSE et les mêmes paramètres que ceux utilisés pour la stratégie avec programmation stochastique (cf. graphique 77). Le même flux de passif sera également considéré dans les deux stratégies. Le nombre de scénarios projetés dans le cas de la stratégie *Fixed-Mix* est égal à 1 000 scénarios. Le logigramme 77 synthétise la démarche générale adoptée. Dans ce schéma, le cercle représente la méthodologie ou le logiciel utilisé. Le rectangle représente, quant à lui, une étape dans l'étude du modèle d'ALM développé.

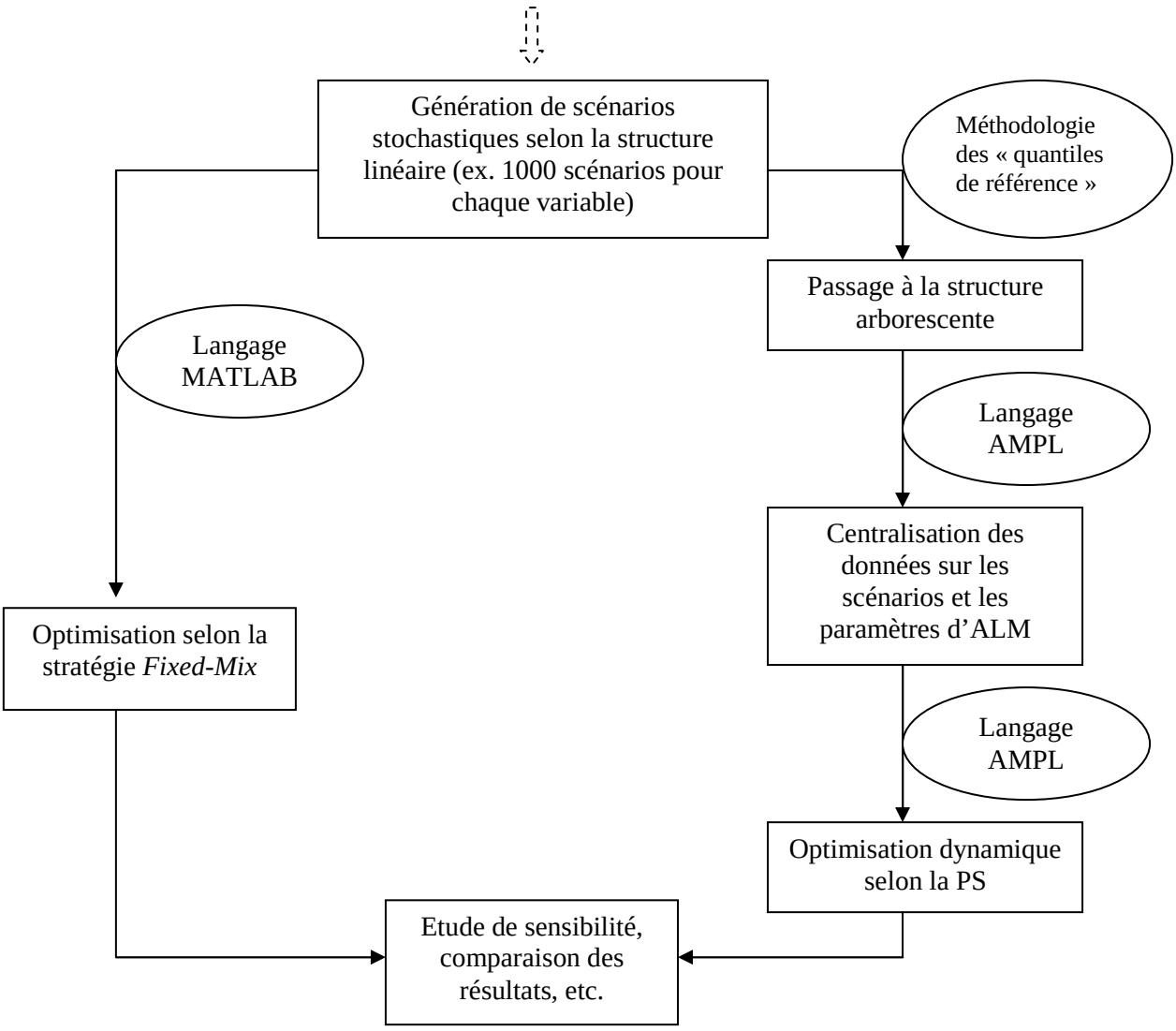


Fig. 77 : *Logigramme des traitements effectués dans le cadre de l'étude du modèle d'ALM développé*

III-3 Discussion des résultats

Nous étudierons ici l'effet du changement des « quantiles de référence » sur l'allocation finale, toutes choses étant égales par ailleurs (cf. graphique 77). De même, la structure des nœuds de cet arbre sera considérée variable : différents tests seront effectués pour mettre en évidence l'impact du changement de cette structure sur les résultats finaux de l'allocation. Un test supplémentaire par rapport au nombre d'années entre deux niveaux de l'arbre est mis en évidence : il s'agit du test de changement de la structure des périodes de l'arbre, toutes choses étant égales par ailleurs.

Le tableau de la structure des nœuds de l'arbre (cf. tableau 40) indique les poids optimaux du portefeuille au niveau $t = 0$ pour différentes structures de nœuds possibles ([1 5 3 3], [1 7 3 3], [1 9 3 3], [1 3 4 4], [1 4 4 4] et [1 5 4 4])[22], toutes choses étant égales par ailleurs.

[22] Rappel : Nous adoptons la convention d'écriture suivante : [1 5 3 3] signifie par exemple que le nœud parent initial a 5 nœuds enfants, lesquels ont chacun 3 nœuds enfants, puis chacun d'entre eux encore 3 nœuds enfants.

Il s'agit de l'allocation initiale recommandée, début 2010, pour chacune de ces structures. Il est supposé que les poids vont être ajustés aux différents niveaux ultérieurs de l'arbre compte tenu des évolutions possibles des variables modélisées. Cela donne lieu en fait non pas à une unique allocation stratégique vers laquelle nous convergeons de façon périodique et systématique via un processus de rebalancement (cas de la stratégie *Fixed-Mix*) mais plutôt à un « chemin » optimal d'allocations avec comme point de départ les allocations du tableau 40. Un exemple du « chemin » optimal d'allocations obtenu sur toute la période de projection (tout au long de l'arbre) est donné dans les graphiques 78 et 79 respectivement pour les deux structures [1 5 3 3] et [1 5 4 4].

Le tableau 40 présente également la valeur de la fonction objectif et la probabilité de ruine estimées pour chacune de ces solutions. Quant à la « réserve initiale minimale », elle correspond à W_0 minimale à partir de laquelle il est possible d'avoir au moins un « chemin » d'allocation résolvant le problème d'optimisation. Nous notons que plus la réserve initiale est élevée, plus la probabilité de ne pas couvrir l'engagement final cible L_s est faible et implicitement l'effet de la pénalité devient négligeable. Le même tableau présente les valeurs de ces différents indicateurs obtenus dans le cas de la stratégie *Fixed-mix*. Il est intéressant d'étudier la convergence entre les deux stratégies lorsque nous augmentons sensiblement le nombre de scénarios mis en jeu dans l'arbre.

Nous constatons que, même si l'allocation initiale recommandée est la même pour les différentes structures (à savoir 100 % d'actions), les valeurs estimées de la fonction objectif et de la probabilité de ruine diffèrent nettement selon la structure des nœuds, en particulier pour le cas des structures qui supposent principalement un découpage en 3 quantiles (à savoir [1 5 3 3], [1 7 3 3], [1 9 3 3]) ainsi que pour celles basées essentiellement sur une projection en 4 quantiles (à savoir [1 3 4 4], [1 4 4 4], [1 5 4 4]). Le premier groupe permet d'avoir des valeurs de la fonction objectif relativement plus élevées avec des niveaux de probabilité de ruine également plus élevés. Autrement dit, pour pouvoir obtenir un meilleur rendement l'investisseur doit aussi assumer plus de risque : nous retrouvons ainsi le cadre de la théorie d'efficience de Markowitz [1952].

Incidemment, nous remarquons que l'effet de la structure des nœuds sur la valeur du couple d'indicateurs (objectif / probabilité de ruine) est très significatif, voire même plus important que l'effet du nombre de scénarios ou du nombre total de nœuds dans l'arbre. En fait, l'augmentation du nombre de scénarios (en passant par exemple de 45 scénarios avec la structure [1 5 3 3] à 80 scénarios avec la structure [1 5 4 4]) ne donne pas forcément lieu à une convergence de la valeur de la fonction objectif vers celle de la stratégie *Fixed-Mix* (31 817 euros), mais au contraire nous constatons la baisse de la valeur de la fonction objectif passant de 28 327 à 24 103, le tout accompagné d'un profil de risque de moins au moins élevé (en particulier, le passage de la structure [1 7 3 3] à la structure [1 5 4 4] induit la baisse du niveau de risque de 19 % à 11 %).

Par ailleurs, plus le nombre de scénarios augmente dans l'arbre, plus le besoin initial en réserve diminue et converge vers celui exprimé dans le modèle *Fixed-Mix*. Par exemple, dans le cas de la structure [1 5 3 3], qui correspond au nombre de scénarios le moins élevé, l'exigence en capital initial est également la plus élevée (8 000). A l'inverse, la structure [1 5 4 4] qui comporte pratiquement le plus grand nombre de scénarios possède, quant à elle, la valeur de la réserve initiale minimale la plus faible. On le voit, le rôle du nombre de scénarios intervient plus au niveau de la réserve initiale minimale exigée qu'au niveau de la convergence vers la solution optimale absolue.

	Structure des nœuds de l'arbre						
	[1 5 3 3]	[1 7 3 3]	[1 9 3 3]	[1 3 4 4]	[1 4 4 4]	[1 5 4 4]	*Fixed Mix*
Monétaire	0%	0	0%	0%	0%	0%	0%
Obligation	0%	0	0%	0%	0%	0%	0%
Action	100%	100%	100%	100%	100%	100%	100%
Fonction objectif	28 327	27 953	28 173	24 093	23 283	24 103	31 817
Probabilité de ruine	13,30%	19,00%	18,50%	12,50%	11,00%	11,25%	13,70%
Réserve initiale minimale	8000	8000	6000	6000	4000	4000	4000
Nombre de scénarios	45	63	81	49	64	80	1000
Nombre total de nœuds	65	91	117	63	84	105	-

Tab. 40 : *Allocations optimales et leurs caractéristiques en fonction des structures de l'arbre retenues et des stratégies retenues (avec la structure des périodes {1 4 15} et les actions comme variable de référence)*

Le tableau 41, relatif à la structure des périodes de l'arbre, montre que les résultats obtenus via la programmation stochastique sont très sensibles au choix de la structure retenue pour les durées des périodes (ou précisément les sous-périodes sur la durée totale de 20 ans), cela en supposant un même flux de passif de départ. Plus précisément, nous procédons en agrégeant sur 20 ans un même flux de passif via différentes structures de période, à savoir les suivantes : {1 4 15}, {1 1 18}, {1 9 10}, {6 7 7}, {10 9 1}, {15 4 1} et {18 1 1}[23].

Nous constatons à ce niveau que les valeurs pour la fonction objectif et pour la probabilité de ruine diffèrent sensiblement selon que l'agrégation est essentiellement effectuée sur les 15 dernières années (par exemple {1 4 15} et {1 1 18}) ou sur les 15 premières années (par exemple {15 4 1} et {18 1 1}). L'exposition au risque action peut également fluctuer de 0 % à 100 % sur l'allocation initiale recommandée. Bien entendu, l'ampleur et le sens de cet impact sur les différents indicateurs (valeur de la fonction objectif, poids optimaux, etc.) ne peut pas être généralisé puisque cela dépend de la série de flux de passif considérée. La réserve initiale minimale nécessaire pour avoir une solution au problème reste, quant à elle, quasi constante, quelle que soit la structure des périodes retenue.

Le tableau 42 étudie la sensibilité de l'allocation optimale par rapport au choix effectué pour la variable de référence (ex. rendement des actions) dans le cadre de la méthodologie proposée (méthodologie « des quantiles de référence »), toutes choses étant égales par ailleurs. L'effet de ce choix sur les valeurs de la fonction objectif, la probabilité de ruine ainsi que la réserve initiale minimale est étudié.

Notons que le changement de la variable de référence a un effet moindre sur l'allocation optimale initiale à t_0 que le changement de la structure des périodes de l'arbre (cf. tableau 41). De même, nous remarquons une baisse significative de la valeur de la fonction objectif, accompagnée par une aversion au risque plus importante lorsque l'on passe des actions comme variable de référence au monétaire : la valeur de la fonction objectif pour la solution optimale passe ainsi de 24 103 dans le cas des actions à 16 587 dans le cas du monétaire tandis que la probabilité de ruine estimée passe respectivement de 11 % à 0 %, indiquant un degré élevé de prudence dans le chemin d'allocations recommandé dans le cas du monétaire.

[23] Rappel : Nous adoptons la convention d'écriture suivante : {1 4 15} signifie par exemple les 3 sous-périodes consécutives 1 ans, 4 ans et 15 ans.

Cela montre l'importance significative du choix de la variable de référence à ce niveau. La réserve initiale minimale nécessaire pour avoir une solution au problème reste, quant à elle, quasi constante quel que soit le quantile de référence retenu.

	Structure des périodes de l'arbre						
	{1 4 15}	{1 1 18}	{1 9 10}	{6 7 7}	{10 9 1}	{15 4 1}	{18 1 1}
Monétaire	0%	0%	83,50%	0,00%	86%	0	0%
Obligation	0%	0%	0%	6,50%	14%	17,5%	15%
Action	100%	100%	16,50%	93,50%	0%	82,50%	85%
Fonction objectif	24 103	24 646	23 459	24 035	18 004	24 476	23 436
Probabilité de ruine	11,25%	15,00%	17,50%	12,50%	5,00%	5,00%	11,00%
Réserve initiale minimale	4000	4000	4000	4000	4000	3000	4000

Tab. 41 : *Allocations optimales et leurs caractéristiques en fonction des horizons retenus (avec la structure des nœuds [1 5 4 4] et les actions comme variable de référence)*

	Variables de référence			
	Action	Monétaire	Obligation	Inflation
Monétaire	0%	32,20%	0%	0%
Obligation	0%	0%	0%	0%
Action	100%	67,80%	100%	100%
Fonction objectif	24 103	16 587	19 467	16 370
Probabilité de ruine	11%	0%	0%	0%
Réserve initiale minimale	4000	3000	4000	4000

Tab. 42 : *Allocations optimales et leurs caractéristiques en fonction des quantiles de référence retenus (avec la structure des nœuds [1 5 4 4] et la structure des périodes {1 4 15})*

Le tableau suivant résume nos différentes constatations sur l'effet du changement de certains facteurs d'entrée (*input*) sur les états de sortie (*output*) du modèle d'ALM développé.

Output / **Facteur**	Fonction objectif /probabilité de ruine	Réserve initiale minimale	Allocation initiale optimale
Structure des nœuds	(+)*	(+)	(-)
Structure des périodes	(+)	(-)	(+)
Variable de référence	(+)	(-)	(+)
Nombre de scénarios	(-)	(+)	(-)

- : (+) si le facteur a un impact significatif sur la valeur de l'indicateur de sortie (*output*)
  (-) si le facteur n'a pas d'impact significatif sur la valeur de l'indicateur de sortie (*output*)

Tab. 43 : *Synthèse de l'impact constaté de certains facteurs sur les résultats obtenus par le modèle d'ALM proposé*

Concernant les deux graphiques 78 et 79, et pour mieux les comprendre, nous décrivons dans un premier temps le graphique 78 (ayant la structure des nœuds [1 5 4 4] et la structure des périodes {1 4 15}). Chaque rectangle dans ce graphique représente l'allocation optimale, selon le modèle proposé d'ALM, à l'instant indiqué et dans le cas du scénario de marché correspondant : c'est pour cette raison que l'arbre des allocations est identique, dans son schéma de présentation, à l'arbre des scénarios. La différence est que l'arbre des allocations

indique, à chaque nœud, l'allocation prévue être optimale plutôt que la performance anticipée (ou projetée) du marché.

Par exemple, l'allocation optimale à l'instant t_0 est indiquée par le vecteur (0 %, 0 %, 100 %) dont les éléments sont les poids respectifs du monétaire, des obligations et des actions. n_{t_0} correspond au nombre de nœuds dans l'arbre à l'instant t_0. En passant à l'instant t_1, autrement dit après une année selon la structure des périodes considérée, nous avons cinq scénarios possibles de marché (n_{t_1} =5). Le modèle d'ALM basé sur la PS indique, pour chacun de ces scénarios, l'allocation optimale correspondante, par exemple (0 %, 100 %, 0 %) pour le scénario de marché le plus défavorable représenté par le rectangle tout à fait en bas dans le graphique. La même lecture peut être appliquée aux instants ultérieurs de prise de décision (notamment t_2) ainsi qu'au graphique 79.

Les graphiques 78 et 79 illustrent deux « chemins » optimaux d'allocations proposées pour une même structure de période (celle du {1 4 15}) et pour deux structures de nœuds différentes (respectivement [1 5 4 4] et [1 5 3 3]). Nous constatons une nette différence entre les deux chemins surtout au niveau des 15 dernières années, ce qui est du au fait que la structure des nœuds diverge sur cette période. La dynamique des poids est mise en évidence via ces deux graphiques : le processus suppose une adaptation « optimale » de la composition du portefeuille compte tenu des scénarios économiques projetés. Les nœuds représentants les « pires » scénarios (ceux qui se situent en bas) tendent souvent à préconiser plus d'obligations que les autres, compte tenu de l'hypothèse que les obligations représentent moins de risque que les actions.

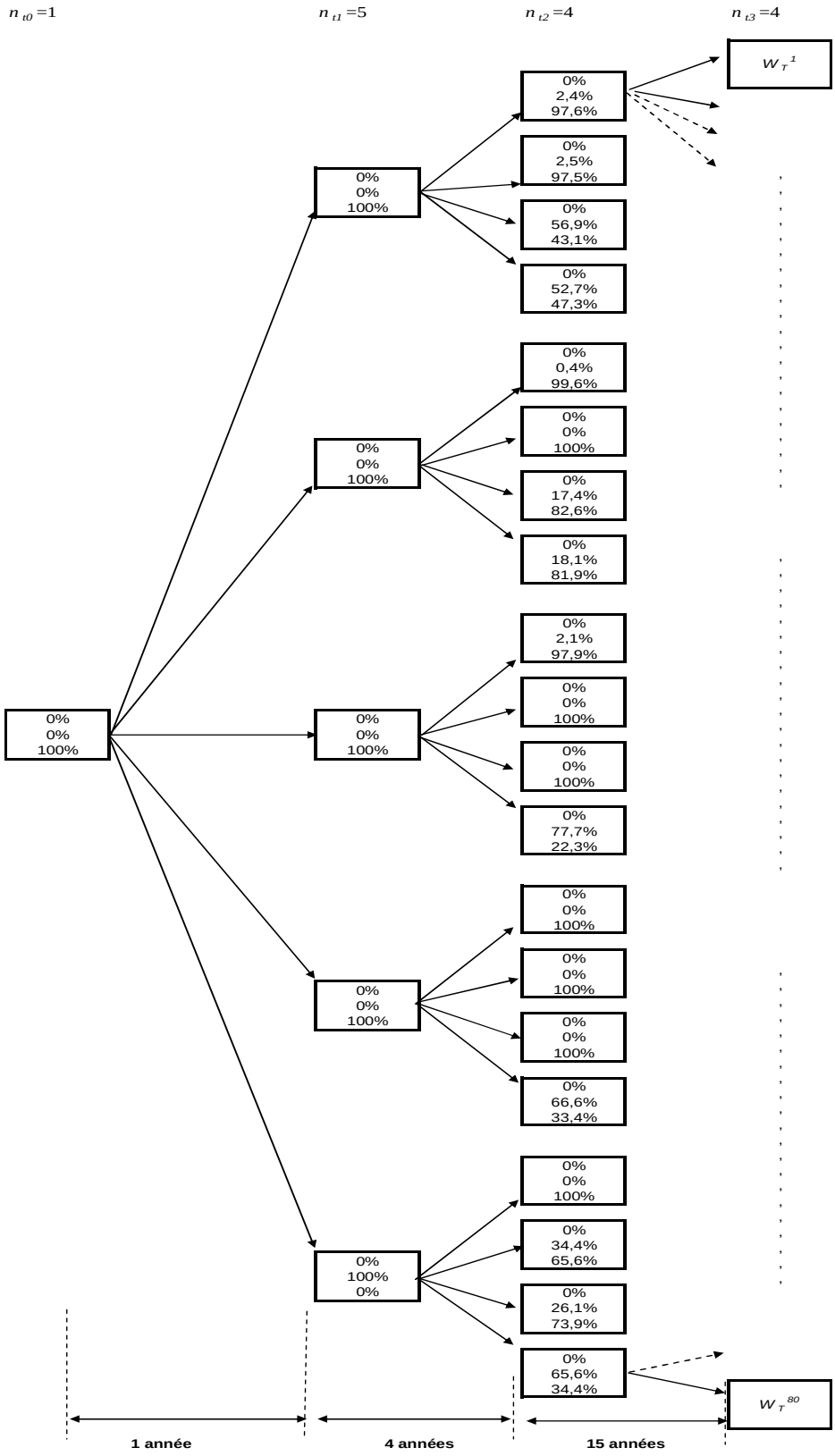


Fig. 78 : *Présentation de l'arbre des allocations (« chemin ») obtenu par le modèle d'optimisation avec PS (cas des paramètres standards, en particulier la structure des nœuds [1 5 4 4] et la structure des périodes {1 4 15})*

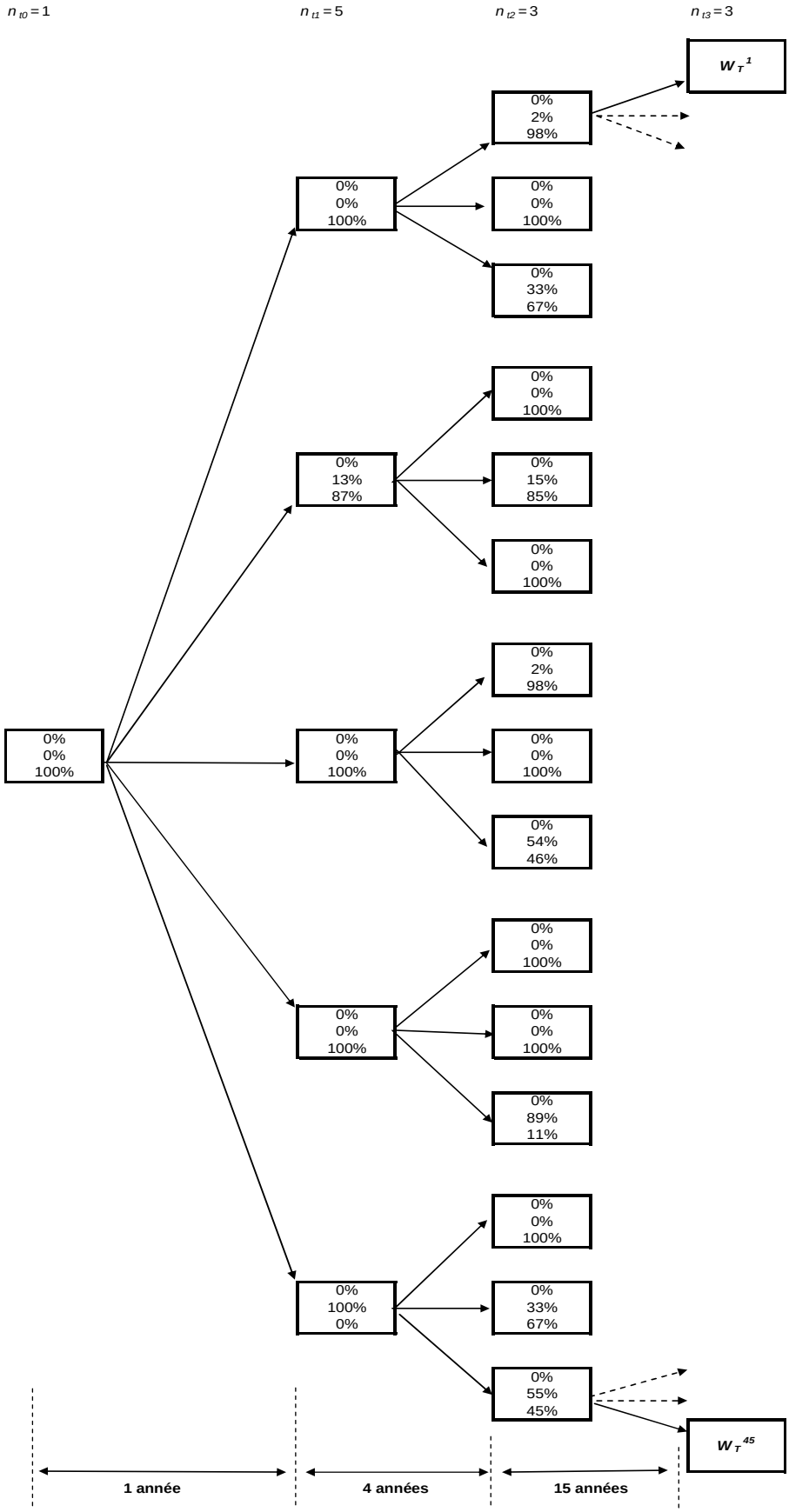


Fig. 79 : *Présentation de l'arbre des allocations (« chemin ») obtenu par le modèle d'optimisation avec PS (cas de paramètres alternatifs, en particulier une structure des nœuds alternative [1 5 3 3] et la structure des périodes {1 4 15})*

Selon le modèle d'allocation stratégique d'actifs développé ici (en se basant sur les techniques de programmation stochastique), la règle générale constatée de cheminement des allocations optimales est la suivante :

- soit la réalisation de la période précédente est favorable (le rendement du portefeuille a augmenté) le portefeuille s'autorise à rajouter de l'actif risqué (notamment les actions),
- sinon le portefeuille tend à se couvrir de plus en plus en actif moins risqué.

A titre d'exemple, dans le graphique 78, le scénario le plus défavorable entre t_0 et t_1 conduit à une allocation « 100 % obligations » (comme le montre le nœud le plus bas à la date t_1) : il s'agit d'une réponse à cette évolution défavorable du rendement du portefeuille avec sa composition précédente (à la date t_0). Les nœuds enfants à t_2 dudit nœud ne préconisent pas la même allocation. En particulier, nous constatons un retour à l'allocation « 100 % actions » pour le nœud enfant qui suit le scénario le plus favorable entre t_1 et t_2. De même, l'allocation reste fortement chargée en obligation suite au scénario le plus défavorable entre ces deux dates (65,6 % obligation contre 34,4 % actions).

De façon plus générale et synthétique, selon notre modèle d'ALM basé sur la programmation stochastique, au fur et à mesure que nous rencontrons des scénarios favorables, l'allocation préconisée commence à retenir de plus en plus d'actions, dans le cas inverse le portefeuille se charge en actifs non risqués. La non utilisation de bornes supérieures pour les poids de différentes classes d'actifs a permis, à notre avis, de mieux mettre en évidence ce constat. En pratique, le recours à l'utilisation de telles bornes est souvent le cas (cf. section I, chapitre 2 de la partie II).

Cela nous emmène à avancer nos deux dernières remarques à ce niveau. Dans la première, nous tenons à rappeler le rôle des valeurs retenues pour les variables d'entrée du modèle : leur modification conduit sûrement à des valeurs de sortie différentes (en particulier pour l'allocation initiale optimale). L'objet a été, rappelons le, d'illustrer et de mettre en évidence le cadre d'utilisation des techniques de PS pour l'allocation stratégique d'un régime de retraire partiellement provisionné. La deuxième remarque a trait à l'élargissement du champ des tests de sensibilités, notamment le test d'impact de structure de nœud plus grande. Cela n'a pas pu être effectué dans cette étude, faute de moyens, mais pourra faire l'objet de prochaines études spécifiques.

Conclusion du chapitre 3

Après avoir discuté du cadre général d'utilisation des programmes stochastiques avec recours, nous avons développé un modèle d'ALM dynamique basé sur ces techniques. Une des spécificités du modèle développé se situe au niveau de la méthodologie de génération de scénarios économiques que nous avons appelée méthodologie « des quantiles de référence ».

Cette dernière permet de partir d'une structure linéaire des scénarios générés (telle que décrite dans le chapitre 2 de la partie I) pour finalement construire l'arbre de scénarios servant de base à l'optimisation avec les techniques de PS. La méthodologie a le mérite de tenir compte de la corrélation entre les distributions des différentes variables projetées et permet de lier l'aspect linéaire de génération des scénarios avec les techniques de PS. En particulier, le modèle d'ALM développé réduit la dimension du problème rencontré avec la stratégie *Fixed-Mix*.

Conclusion de la partie II

Face aux limites de la stratégie *Fixed-Mix*, une stratégie d'ALM dynamique basée sur la programmation stochastique a été développée dans cette partie. Elle pourra servir comme une approche de référence, dont les résultats sont à comparer avec ceux de la stratégie *Fixed-Mix*. Cette stratégie peut également constituer un outil pour tester la fiabilité et l'optimalité de l'allocation de long terme obtenue avec la stratégie *Fixed-Mix*. Une limite majeure de la programmation stochastique, à l'opposée de la programmation dynamique par exemple, est que les distributions de probabilité des paramètres aléatoires sont supposés être données et ne peuvent pas dépendre des décisions prises (cf. Zenios [2007]).

Les techniques d'optimisation numérique ont été également l'un des sujets développés dans la partie II. L'algorithme proposé a pour but de rechercher des paramètres de $\Re^d$ conduisant à minimiser, globalement, une fonction réelle bruitée. L'algorithme permet de construire des zones de confiance autour des paramètres solutions supposés. Il permet également de moduler, au moyen des deux paramètres σ_K et α, la réitération de tirages dans un voisinage des points solutions découverts, ou à l'inverse l'exploration de la fonction dans des zones encore peu explorées.

Comme nous l'avons vu, les sujets de réflexions ne manquent pas pour développer des modèles plus robustes et mieux à même de rendre compte de la complexité des risques économiques et financiers supportés par un fonds de retraite. L'étude menée dans cette partie s'insère dans le cadre de la recherche d'une vision à la fois simplifiée, réelle et dynamique des approches possibles pour l'allocation stratégique d'actifs d'un régime de retraite partiellement provisionné.

CONCLUSION GÉNÉRALE

Synthèse et analyse des résultats

L'étude de l'utilité des réserves pour un système par répartition et *a fortiori* de leur gestion reste un sujet peu exploré. Pour apprécier la solidité d'un régime de retraite, les experts se sont longtemps limités à une approche binaire, selon laquelle il convenait de tout provisionner dans le cas d'un système par capitalisation et rien dans le cas d'un système par répartition. C'est dans ce contexte que le présent travail a été focalisé sur les modèles d'allocation stratégiques d'actifs et sur leurs applications pour la gestion des réserves financières des régimes de retraite par répartition.

Au terme de cette étude, nous pouvons tenter de tirer un certain nombre de conclusions et de points clés des travaux présentés :

- Concernant les générateurs de scénarios économiques (GSE) :

    - Le choix des composantes d'un GSE est lié à sa vocation finale, que ce soit en tant qu'outil d'évaluation des produits financiers (*pricing*) ou en tant qu'outil de projection et de gestion des risques.

    - Dans le cas où le GSE est utilisé comme un outil de gestion des risques, il ne faut pas perdre de vue qu'il ne s'agit finalement que d'un outil d'aide à la prise de décision. De ce point de vue, la performance du GSE ne peut être mesurée que sur la base de sa contribution à la prise de la « bonne » décision et non sur la base du degré de correspondance des scénarios projetés par rapport à la distribution réelle des variables du GSE, une distribution qui reste peu accessible quel que soit les outils d'estimation et d'approximation employés. A ce niveau, certains indicateurs de mesure de performance ont été étudiés, à savoir : l'indicateur de stabilité et l'indicateur d'absence de biais.

    - La distinction entre les aspects théoriques et la mise en œuvre lors de l'élaboration d'un GSE est nécessaire. Par aspects théoriques, nous notons en particulier la conception et le choix de la structure de dépendance. Par mise en œuvre nous notons en particulier la détermination de la structure de projection des scénarios, le choix de la méthodologie de leur génération et le calibrage du générateur.

    - Au-delà de la problématique des queues de distribution des variables modélisées, l'instabilité dans le temps de la corrélation entre les différents facteurs de risque a été considéré comme un élément important lors du choix du niveau de finesse d'un GSE. Cette importance provient de l'impact significatif des hypothèses de corrélation sur les résultats finaux obtenus et ainsi sur la décision finale de l'investisseur (cas de l'allocation d'actifs). Par exemple, l'omission des régimes qui caractérisent les processus générateurs de la rentabilité des actions dans l'analyse de la relation « inflation-rentabilité peut introduire des biais.

      Nous avons mis en évidence la nécessité de ne pas sous estimer l'impact de l'instabilité dans le temps des corrélations et ainsi l'importance de ne pas se contenter des modèles log-normaux usuels et/ou de corrélations linéaires. A

ce niveau, nous avons essayé, en construisant un GSE, d'apporter des améliorations par rapport à ce qui se pratique usuellement en proposant un modèle de changement de régime plus global par rapport au modèle initial de Hardy [2001].

En particulier, la notion de changement de régime a été élargie pour concerner non seulement les rendements et les volatilités des actions, comme dans le cas du modèle initial de Hardy [2001], mais aussi la matrice des corrélations entre les différentes variables. Cette matrice devient fonction du régime de marché observé à une date donnée : cela nous place dans un cadre de dépendance non linéaire. L'apport de cette approche au niveau de l'allocation stratégique d'actifs a été mis en évidence via une application numérique. Les résultats obtenus confirment une aversion au risque supplémentaire par l'investisseur lors de l'utilisation de telles approches.

- Dans le cadre de notre construction d'un GSE, il apparaît délicat de prétendre obtenir des hypothèses pérennes sur les paramètres de celui-ci et de retenir une approche de calibrage plutôt qu'une autre de manière définitive. Le processus de calibrage dépend fortement de différents facteurs (nature de données, historique retenu, etc.) et les résultats obtenus par le GSE sont dépendent, bien évidemment, des hypothèses initiales sur les différents paramètres. Cela ne peut être remédié *a priori* que par un recours à des *stress test* sur les hypothèses initiales. Dans le cadre du calibrage du GSE construit, un ensemble d'outils a été présenté de manière à permettre l'estimation de ses différents paramètres.

▪ Concernant les modèles d'élaboration de l'allocation d'actif

- L'aspect mono-périodique qui caractérise les différents modèles classiques déterministes de gestion actif-passif (que ce soit celles basées sur les techniques d'immunisation ou celles basées sur la notion du surplus) limite leur capacité à refléter de façon fiable les perspectives d'évolution des différentes variables financières, en particulier sur le long terme. Cela devient d'autant plus compliqué que les variables à considérer sont plus nombreuses et que la nécessité de prendre en compte les dépendances entre les différentes variables est plus importante.

- Lors du choix des critères de l'allocation stratégique, il est nécessaire de tenir compte du type de régime de retraite considéré (régime provisionné, régime partiellement provisionné, régime non provisionné). En outre, quel que soit le régime, dans le développement de l'allocation stratégique, une attention particulière peut également être portée à la distinction entre la phase de constitution et la phase de restitution du régime.

- En guise d'illustration, la sensibilité des résultats de l'allocation d'actifs par rapport aux hypothèses retenues initialement sur les paramètres du GSE a été mise en évidence dans le cadre de la stratégie *Fixed-Mix*.

- L'approche dynamique pour l'allocation stratégique d'actifs présente l'avantage théorique de la robustesse face aux changements de régime des

marchés. L'autorisation du changement des poids des différentes classes d'actifs, sur la base d'une règle de gestion bien définie, constitue un élément intéressant. Au-delà du fait que cela permet l'ajustement des expositions aux différentes classes d'actifs (suite à l'évolution des conditions de marché), cette approche peut constituer dans certaine mesure une version simplifiée et réaliste de l'approche *Fixed-Mix* (à poids constants). Certains de ces modèles dynamiques ont été mis en œuvre et détaillées (notamment, les techniques d'assurance de portefeuille)

- Parmi les modèles dynamiques d'allocation d'actifs, les techniques de « programmation stochastique » (cf. Dantzig et al. [1990]) sont peu explorées, en particulier au niveau des régimes de retraite par répartition partiellement provisionnés. Nous sommes ainsi partis de la description du principe de base de ces techniques pour finalement mettre en œuvre des techniques novatrices d'ALM.

  Parallèlement, une nouvelle méthodologie pour la génération de l'arbre des scénarios a été adoptée. La méthodologie a le mérite de tenir compte de la corrélation entre les distributions des différentes variables projetées et permet de lier l'aspect linéaire de génération des scénarios (tel que décrit dans la partie I de ce travail) avec les techniques de PS (sujet profondément étudié dans la partie II).

  Une étude comparative du modèle d'ALM ainsi développé avec celui basé sur la stratégie *Fixed-Mix* a été effectuée. L'écart constaté en termes de résultats peut *a priori* être expliqué par les paramètres d'entrée de chacun des deux modèles, qui ne permettent pas forcément de converger vers un même résultat.

  Différents tests de sensibilité ont été par ailleurs mis en place pour mesurer l'impact du changement de certaines variables clés d'entrée sur les résultats produits par notre modèle d'ALM. Les résultats obtenus, dans le cadre de ces tests, indiquent une sensibilité significative à certains paramètres ce qui signifie qu'une attention particulière doit être donnée aux hypothèses de départ.

  Concernant la règle de gestion qu'adopte, implicitement, notre modèle d'ALM basé sur la programmation stochastique, nous avons constaté qu'au fur et à mesure que nous rencontrons des scénarios favorables, l'allocation optimale préconisée par le modèle commence à retenir de plus en plus d'actifs risqués (ex. les actions selon nos hypothèses), dans le cas inverse le portefeuille se charge en actifs non risqués.

De façon générale, la programmation stochastique est un domaine en pleine effervescence, permettant d'étendre l'optimisation classique au niveau des problèmes stochastiques et dynamiques tout en gardant la possibilité de résoudre ce type de problème. L'approche de programmation stochastique développée dans ce travail est claire, théoriquement efficiente et relativement facile à mettre en œuvre (implémentation).

L'application de cette approche peut améliorer les techniques actuelles utilisées par les fonds de retraite en France en matière d'allocation stratégique d'actifs, souvent basées sur la simple simulation et/ou sur l'expérience. Le modèle d'ALM développé (en se basant sur les techniques de PS) pourra également servir comme une approche de référence, dont les résultats sont à comparer avec ceux de la stratégie *Fixed-Mix*. Cette stratégie peut aussi constituer un outil pour tester la fiabilité et l'optimalité de l'allocation de long terme obtenue avec la stratégie *Fixed-Mix*.

Cette thèse a également accordé une attention particulière aux techniques numériques de recherche de l'optimum, qui demeurent des questions essentielles pour la mise en place d'un modèle d'allocation. Le point de départ était le constat d'un temps de calcul significatif dû simultanément à un nombre élevé de scénarios économiques générés et à un nombre d'allocations d'actifs testées également élevé.

Cela nous a conduits à une réflexion sur un algorithme d'optimisation globale d'une fonction non convexe et bruitée. L'algorithme est construit après une étude de critères de compromis entre, d'une part, l'exploration de la fonction objectif en de nouveaux points (correspondant à des tests sur de nouvelles allocations d'actifs) et d'autre part l'amélioration de la connaissance de celle-ci, par l'augmentation du nombre de tirages en des points déjà explorés (correspondant à la génération de scénarios économiques supplémentaires pour les allocations d'actifs déjà testées). Une application numérique a illustré la conformité du comportement de cet algorithme à celui prévu théoriquement.

L'algorithme permet de moduler facilement, au moyen de deux paramètres, la réitération de tirages dans un voisinage des points solutions découverts, ou à l'inverse l'exploration de la fonction dans des zones encore peu explorées. En outre, une procédure d'estimation de la régularité fonction objectif f permet d'ajuster les valeurs de ces deux paramètres. Il s'agit d'un algorithme qui vise à montrer l'intérêt et la faisabilité technique d'une grille à pas variable, susceptible de répondre au problème du choix entre exploration et connaissance d'une fonction aléatoire.

Perspectives

Les réserves contribuent à part entière à la solidité du régime. L'ensemble des éléments présentés fournit des composants essentiels pour élaborer l'allocation stratégique d'un régime de retraite partiellement provisionné. Une première perspective de prolongement de ce travail pourra être d'examiner l'apport de chacune des classes d'actifs dites « non classiques » (tels que l'immobilier, les obligations indexées sur l'inflation, etc.) en tant qu'alternative stratégique d'investissement. Nous rappelons que dans le cas de l'allocation stratégique d'actifs d'un fonds de retraite la priorité est souvent donnée aux actions, aux obligations et au monétaire, dites classes « classiques » (cf. Campbell et al. [2001]).

En matière de projection de grandeurs réelles sur le long terme, les actifs non classiques sont souvent traités avec prudence et ne suscitent pas la grande part de l'intérêt des décideurs. Ces classes d'actifs se heurtent souvent aux problèmes d'un historique peu profond, d'une liquidité insuffisante et de données confidentielles (cas des fonds de couverture). Cela ne remet pas en cause le potentiel qu'elles présentent en tant que source de performance et/ou de couverture supplémentaire pour les réserves gérées : dans ce sens leurs apports à la gestion des réserves pourront faire l'objet d'études spécifiques sur le court terme. La littérature

présente à ce niveau différentes techniques pouvant servir de base à cette étude en particulier les techniques d'Analyse par Composante Principale (cf. Roncalli [1998]).

Par ailleurs, nous notons qu'un régime de retraite par répartition provisionné (ou partiellement provisionné) doit faire face au moins à deux risques essentiels : l'incertitude sur l'évolution future de ses flux techniques et le risque de placement. Le présent travail a été focalisé sur la gestion et le contrôle du deuxième risque. Il apparaît ainsi intéressant la réalisation d'analyses de sensibilité à la tendance des flux techniques futurs (risque de longévité, etc.).

Cela nous place dans un cadre d'étude plus global permettant de définir les termes d'un pilotage technique du régime. Par pilotage technique on entend les règles d'actualisation, au fil du temps, des paramètres du régime notamment les taux de cotisation et de prestation en fonction de l'évolution de la solidité financière du régime. L'étude des possibilités d'application des techniques de programmation stochastique avec recours pour ce type de problématique constitue *a priori* un sujet prometteur.

Enfin, nous travaillons à enrichir l'algorithme d'optimisation numérique par exploration sélective. Le cadre de son application a été jusque là réduit à un cas simplifié pour des raisons pratiques de mise en œuvre. L'un des prochains défis sera donc de l'exploiter dans un univers plus réaliste et plus global.

Index des graphiques

Fig. 1 : Pourcentages de la performance globale expliqués par certaines composantes de l'allocation d'actifs selon l'étude de Brinson et al. [1991].....……8
Fig. 2 : Les modules de la formule standard du SCR……...13
Fig. 3 : Structure par cascade dans le modèle de Wilkie [1986]……...26
Fig. 4 : Structure du modèle d'Ahlgrim et al. [2005]……28
Fig. 5 : Comparaison entre deux structures schématiques de projection des scénarios……31
Fig. 6 : Exemple détaillé de la structure d'arbre des scénarios…….31
Fig. 7 : Arbre de scénarios avec différents nombres de nœuds-enfants pour les nœuds d'un même niveau…….32
Fig. 8 : Projection du rendement des actions selon une structure schématique linéaire……...47
Fig. 9 : *Box plot* de la distribution des valeurs de la fonction objectif pour différentes tailles de scénarios……48
Fig. 10 : Les allocations optimales obtenues pour les différents ensembles de scénarios (30 ensembles) ayant chacun la taille de 10 000 scénarios…….49
Fig. 11 : Evolution de la corrélation (glissante) sur 5 ans entre l'inflation et le rendement des actions aux Etats-Unis depuis 1970…….56
Fig. 12 : La corrélation entre les taux d'inflation et les taux nominaux courts au Royaume-Uni……59
Fig. 13 : Illustration du phénomène de retour à la moyenne des taux réels dans le cadre du modèle de Hull et White à deux facteurs……..60
Fig. 14 : Illustration de la structure du GSE construit……..72
Fig. 15 : Illustration de la structure schématique linéaire dans le cas de la projection des rendements des actions…….93
Fig. 16 : Illustration du passage de browniens indépendants à des browniens corrélés négativement…….95
Fig. 17 : Courbe des taux des obligations indexées sur l'inflation d'une maturité de quinze ans……...100
Fig. 18 : Structure par terme des taux nominaux, des taux réels et des taux d'inflation en cas de taux réels longs élevés par rapport à ceux de l'équilibre : Mise en évidence de la courbure de la STTI……101
Fig. 19 : Evolution dans le temps de la STTI : Mise en évidence du phénomène de retour à la moyenne……..102
Fig. 20 : Evolution des pentes moyennes de taux (écarts absolus moyens entre les taux de deux maturités différentes) tout au long de la période de projection……..103
Fig. 21 : Histogramme des taux nominaux de court terme (3 mois) à la fin de la première période…….104
Fig. 22 : Histogramme des taux nominaux de long terme (10 ans) à la fin de la première période…….104
Fig. 23 : Histogramme des taux réels sur le court terme (3 mois) à la fin de la première période…….104
Fig. 24 : Histogramme des taux réels de long terme (10 ans) à la fin de la première période…….104
Fig. 25 : Histogramme des taux d'inflation anticpés de court terme (3 mois) à la fin de la première période…….104
Fig. 26 : Histogramme des taux d'inflation anticpés de long terme (10 ans) à la fin de la première période…….104
Fig. 27 : Histogramme de la distribution des obligations nominales gouvernementales de duration 5 ans sur la première période……105

Fig. 28 : Histogramme de la distribution des obligations indexées inflation de duration 8 ans sur la première période……105
Fig. 29 : Histogramme de la distribution des rendements des obligations nominales non gouvernementales de duration 5 ans sur la première période……105
Fig. 30 : Evolution du marché des actions entre le régime 1 (actions à faible volatilité) et le régime 2 (actions à forte volatilité) sur une trajectoire simulée……..108
Fig. 31 : Frontière efficiente (Kim&Santomero)……127
Fig. 32 : Frontière efficiente et contrainte de déficit (Kim&Santomero)……...127
Fig. 33 : Frontière efficiente et contrainte de déficit (Sharpe&Tint)……..130
Fig. 34 : Exemple de l'impact de la variation du seuil de rentabilité de l'actif……..133
Fig. 35 : Exemple de l'impact de la variation de la probabilité de déficit de l'actif……..133
Fig. 36 : Exemple de l'impact de la variation du seuil de rentabilité du surplus……134
Fig. 37 : Exemple de l'impact de la variation du ratio de financement initial……134
Fig. 38 : Représentation graphique des portefeuilles vérifiant la contrainte sur la rentabilité relative……136
Fig. 39 : Exemple de la variation de la duration du surplus en fonction de l'écart entre la duration de l'actif et celle du passif……138
Fig. 40 : Illustration du modèle de gestion actif-passif dans le cadre stochastique……142
Fig. 41 : Exemple d'évolution des flux nets de trésorerie (flux désinflatés)……..158
Fig. 42 : Illustration de la structure schématique linéaire dans le cas de la projection des rendements des actions……159
Fig. 43 : Illustration du modèle de gestion actif-passif stochastique objet de l'application dans le cadre de la stratégie *Fixed-Mix*……...160
Fig. 44 : Ratio de couverture des flux jusqu'en 2029 sur un horizon de 9 ans (2018) et un niveau de confiance de 97,5 %.....164
Fig. 45 : Sensibilité de l'allocation optimale par rapport à l'espérance de rendement des actions……165
Fig. 46 : Distribution de la valeur de la réserve totale (nette des flux de trésorerie) dans le cas de projection de scénarios avec deux matrices de corrélation……167
Fig. 47 : Un exemple de partition d'une zone dans $\Re^2$ en 60 zones. (d = 2)……175
Fig. 48 : Allure générale de la fonction test $f(\theta)$ pour $\theta \in Z_0$ ……..192
Fig. 49 : Ensemble des points de simulations de F pour une variabilité σ_K = 2 et un niveau de bruit σ_B = 0 (scission systématique)……...193
Fig. 50 : Ensemble des points de simulations de F pour une variabilité σ_K = 2 et un niveau de bruit σ_B = 0.1 (scission systématique), $\vec{\rho}$ = (0.10, 3.5E-03, 6.8E-07, 1.9E-2)……195
Fig. 51 : Ensemble des points de simulations de F pour une variabilité σ_K = 2 et un niveau de bruit σ_B = 0.3 (scission systématique), $\vec{\rho}$ = (0.13, 5.9E-03, 1.6E-06, 5.8E-2)……196
Fig. 52 : Ensemble des points de simulations de F pour une variabilité σ_K = 20 et un niveau de bruit σ_B = 0.1 (scission systématique), $\vec{\rho}$ = (0.45, 2.2E-03, 7.3E-05, 0.72)……196
Fig. 53 : Ensemble des points de simulations de F pour une variabilité σ_K = 100 et un niveau de bruit σ_B = 0.1 (scission systématique), $\vec{\rho}$ = (0.60, 8.8E-03, 5.9E-04, 0.97). La recherche de l'optimum d'une fonction constante bruitée conduit à un nuage de points d'allure proche de celui-ci……197
Fig. 54 : Ensemble des points de simulation de F pour un niveau de bruit σ_B = 0.1, critère de scission n° 2, pour une variabilité (σ_K ; α) = (1, 1.5), et un nombre minimal de tirage n_0 = 10. La surface des bulles est proportionnelle au nombre de tirages de F en chaque point. $\vec{\rho}$ = (0.085, 8.8E-03, 3.1E-05, 8.2E-04)……197

Fig. 55 : Ensemble des points de simulation de F pour un niveau de bruit σ_B = 0.2, critère de scission n° 2, pour une variabilité (σ_K , α) = (1, 1.5), et un nombre minimal de tirage n_0 = 10. La surface des bulles est proportionnelle au nombre de tirages de F en chaque point. $\vec{\rho}$ = (0.10, 8.8E-03, 6.7E-05, 6.2E-03)……...198
Fig. 56 : Ensemble des points de simulation de F pour un niveau de bruit σ_B = 0.2, critère de scission no 1, pour une variabilité (σ_K , α) = (1, 1.5), et un nombre minimal de tirage n_0 = 10. La surface des bulles est proportionnelle au nombre de tirages de F en chaque point. $\vec{\rho}$ = (0.10, 0.011, 1.5E-05, 1.6E-02)……..198
Fig. 57 : Points d'explorations obtenus, lorsque σ_B = 0.1, avec les paramètres estimés hors bruit, (σ_K , α)= (1.0743, 1.5665) (à gauche) ou estimés en présence de bruit (σ_K , α) = (0.5545; 1.2209) (à droite)……..202
Fig. 58 : Ecart-type empirique de $IR_\alpha(Z)$ obtenu pour la zone Z_0 correspondant au simplexe orthogonal standard initial (en pointillé) et pour la demi-zone Z_1 correspondant à la partie de Z_0 située au dessus de la première bissectrice (trait plein)……202
Fig. 59 : Distribution *a priori* de IR_α (courbe gaussienne continue au premier plan) et histogramme a posteriori, pour la demi-zone Z_1 obtenue après la première itération de l'algorithme. Cas α = 1.2 (à gauche) et α = 1.5 (à droite). Les écarts-type empiriques obtenus à partir de l'histogramme sont respectivement σ = 0.41 et σ = 0.73……203
Fig. 60 : Points d'explorations obtenus pour l'algorithme à scission systématique, cas a, lorsque σ_B = 0.1, avec les paramètres (σ_K , α) = (0.6, 1.2) (à gauche). Zone finalement retenue $S_{m^*,5\%}$ avec un seuil de 5 % (à droite)…….205
Fig. 61 : Points d'explorations obtenus pour l'algorithme avec retirage, cas b, critère de scission n° 1 de comparaison des erreurs de grille et d'estimation, lorsque σ_B = 0.1, avec les paramètres (σ_K , α)= (0.6, 1.2) (à gauche). Zone finalement retenue $S_{m^*,5\%}$ avec un seuil de 5 % (à droite)……..206
Fig. 62 : Points d'explorations obtenus pour l'algorithme avec retirage, cas c, lorsque σ_B = 0.1, critère de scission n° 2 considérant l'ajout de n_0 observations, avec les paramètres (σ_K , α)= (0.6, 1.2) (à gauche). Zone finalement retenue $S_{m^*,5\%}$ avec un seuil de 5 % (à droite)……..206
Fig. 63 : Points d'explorations obtenus pour l'algorithme Kiefer-Wolfowitz-Blum, pour 2000 tirages avec $\sigma_B = 0.1/\sqrt{10}$ (à gauche), ou 20000 tirages avec σ_B = 0.1 (à droite)……...211
Fig. 64 : Ratio de Sharpe du surplus final en fonction de *m*…….……..220
Fig. 65 : VaR à 95 % du surplus final en fonction de *m*…….221
Fig. 66 : Surplus final moyen en fonction de *m*……..221
Fig. 67 : Probabilité de surplus négatif en cours de vie du fonds (en fonction de *m*)…….221
Fig. 68 : Probabilité de surplus final négatif en fonction de *m*……...222
Fig. 69 : Trajectoire moyenne du portefeuille (actif et garantie) pour *m*=2.5……222
Fig. 70 : Allure des 1000 trajectoires simulées du surplus (*m*=2.5)……..….222
Fig. 71 : Illustration du problème de planification à deux étapes (dans le cas d'une seule période)…...……231
Fig. 72 : Illustration du problème de planification à trois étapes (dans le cas de deux périodes)……..233
Fig. 73 : Exemple de la structure schématique d'arborescence pour les trajectoires simulées…...……243

Fig.74 : Illustration du lien entre le niveau de probabilité cumulée α_i d'un quantile q_i et le poids associé à ce dernier dans l'arbre des scénarios (cas où n =3)……...…246
Fig.75 : Illustration du lien entre le niveau de probabilité cumulée α_i d'un quantile q_i et le poids associé à ce dernier dans l'arbre des scénarios (cas où n =3 et α_i est régulier) : $p_1 = p_2 = p_3$……...…247
Fig. 76 : Flux de passif (désinflatés) sur 20 ans……..251
Fig. 77 : Logigramme des traitements effectués dans le cadre de l'étude du modèle d'ALM développé……252
Fig. 78 : Présentation de l'arbre des allocations (« chemin ») obtenu par le modèle d'optimisation avec PS (cas des paramètres standards, en particulier la structure des nœuds [1 5 4 4] et la structure des périodes {1 4 15})……257
Fig. 79 : Présentation de l'arbre des allocations (« chemin ») obtenu par le modèle d'optimisation avec PS (cas de paramètres alternatifs, en particulier une structure des nœuds alternative [1 5 3 3] et la structure des périodes {1 4 15})…….258

Bibilographie

Aarts E.H.L., Laarhoven V. [1985] « Statistical cooling : a general approach to combinatorial optimization problems », *Philips Journal of Research*, 40 [4], 193-226.

Adam A. [2007] *Handbook of Asset and Liability Management: From Models to Optimal Return Strategies*, Wiley, Octobre 2007, 576 pages.

Ahlgrim K.C., D'Arcy S.P., Gorvett R.W. [2005] « Modeling Financial Scenarios: A Framework for the Actuarial Profession », *Proceedings of the Casualty Actuarial Society*, 177-238. http://www.casact.org/pubs/proceed/proceed05/05187.pdf

Ahlgrim K.C., D'Arcy S.P., Gorvett R.W. [2008] « A Comparison of Actuarial Financial Scenario Generators », *Variance*, 2:1, 2008, pp. 111-134.

Albeanu G., Ghica M., Popentiu-Vladicescu F. [2008] « On using bootstrap scenario generation for multi-period stochastic programming applications », *Proceedings of ICCCC 2008*, 15-17 May 2008 , pp. 156-161.

Alexander C. [2001] *Market Models: A Guide to Financial Data Analysis*, Chichester, England: John Wiley & Sons.

Alexander C., Leigh C. [1997] « On the covariance matrices used in Value-at-Risk models ». *Journal of Derivatives*, 4, 50-62.

Alliot J.M. [1996] *Techniques d'optimisation stochastique appliquées aux problèmes du contrôle aérien*, INPT, Habilitation à Diriger des Recherches.

Arbulu P. [1998] *Le marché parisien des actions au XIXe siècle: performance et efficience d'un marché émergent*, Thèse de doctorat ès sciences de gestion, Université d'Orléans.

Armel K., Planchet F., Kamega A. [2010] « Quelle structure de dépendance pour un générateur de scénarios économiques en assurance ? », *Les cahiers de recherche de l'ISFA*, WP2134.

Artus P., [1998] « Faut-il introduire les prix d'actifs dans la fonction de réaction des banques centrales ? », *Document de travail*, Caisse des Dépôts et Consignations, n°1998-26, juin.

Bassetto M. [2004] « Negative Nominal Interest Rates », *American Economic Review (Papers and Proceedings)* 94.2 (2004): 104-108.

Beasley J.E. [1990] « OR-Library: distributing test problems by electronic mail », *Journal of the Operational Research Society*, 41(11) (1990) pp1069-1072. http://people.brunel.ac.uk/~mastjjb/jeb/info.html

Beder T. [1995] « VaR: Seductive but Dangerous », *Financial Analysts Journal*, 51, 12-24.

Bellman R.E. [1957] *Dynamic Programming*, Princeton University Press, Princeton, NJ.

Bhansali V., Wise M. [2001] « Forecasting portfolio risk in normal and stressed markets », *Journal of Risk*, 4, 91-106.

Bienvenüe A., Rullière D. [2010] « Iterative adjustment of survival functions by composed probability distortions », *soumis*.

Birge J.R., Louveaux F. [1997] *Introduction to Stochastic Programming*, Springer Series in Operations Research and Financial Engineering Heidelberg, 1997.

Black F., Litterman R. [1992] « Global Portfolio Optimization », *Financial Analysts Journal*, September/October, 28-43.

Black F., Scholes M. [1973] « The pricing of options and corporate liabilities », *Journal of Political Economy*, 1973, pp. 637-654.

Blake D., Lehmann B., Timmermann A. [1999] « Asset Allocation Dynamics and Pension Fund Performance » , *Journal of Business*, 72, 429-461.

Blanke D. [2002] « Estimation du coefficient de régularité locale d'une trajectoire de processus », *Comptes rendus de l'Académie des sciences*, Paris, Serie I, 334, 145-148.

Blum J.R. [1954] « Multidimensional stochastic approximation methods », *Annals of Mathematical Statistics*, 25, 737-744.

Bluhm C., Overbeck L. [2007] « Calibration of PD term structures: to be Markov or not to be », *Risk*, 20(11), pages 98-103.

Bollerslev T. [1986] « Generalized autoregressive conditional heteroscedasticity », *Journal of Econometrics*, 31, 307-327.

Bouroche J.M., Saporta G., [1980] *L'Analyse des Données*, collection Que sais-je, PUF.

Brandimarte P. [2006] *Numerical Methods in Finance and Economics: A MATLAB-Based Introduction*, Wiley-Interscience, deuxième édition.

Branke J., Meisel S., Schmidt C. [2008] « Simulated annealing in the presence of noise », *Journal of Heuristics*, vol. 14, no 6, pp.627-654.

Brennan M., Schwartz E. [1982] « An Equilibrium model of Bond Pricing and a Test of Market Efficiency », *Journal of Financial and Quantitative Analysis*, 17, 3, 301-329.

Brinson G. P., Singer B.D., Beebower G.L. [1991] « Determinants of portfolio performance II : An update », *Financial Analysts Journal*, May-June 1991.

Broadie M., Cicek D.M., Zeevi A. [2009] « An adaptative multidimensional version of the kiefer-wolfowitz stochastic approximation algorithm », *Proceeding of the 2009 Winter Simulation Conference*.

Bulger D.W., Romeijn H.E. [2005] « Optimising noisy objective functions », *Journal of Global Optimization*, 31 : 599-600.

Byström H.N. [2002] « Using simulated currency rainbow options to evaluate covariance matrix forecasts », *Journal of International Financial Markets, Institutions and Money*, 12, 216-230.

Campbell J.Y., Viceira L.M. [2001] *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press.

Campbell J.Y., Viceira L.M. [2005] « The term structure of the risk-return trade off », *Financial Analysts Journal*, 61, 34-44.

Cariño D.R., Kent T., Myers D.H., Stacy C., Sylvanus M., Turner A., Watanabe K., Ziemba W. T. [1994] « The Russell-Yasuda Kasai Model : An Asset Liability Model for a Japanese Insurance Company using Multi-stage Stochastic Programming », *Interfaces*, 24, Jan-Feb 1994,29-49.

Cariño, D.R., Turner A. [1998] « Multiperiod Asset Allocation with Derivative Assets », dans Worldwide Asset and Liability Modelling, Ziemba & Mulvey, Cambridge University Press, Cambridge , UK, pp.182-204.

Carter J. [1991] « The derivation and application of an Australian stochastic investment model », *Transactions of the Institute of Actuaries of Australia*. 1991, I 315-428.

Castro J. [2009], « A stochastic programming approach to cash management in banking », *European Journal of Operational Research*, 2009, vol. 192, issue 3, pages 963-974.

Chan L.K., Karceski J., Lakonishok J. [1999] « On portfolio optimization: forecasting covariances and choosing the risk model », *Review of Financial Studies*, 12, 937-974.

Chan K.C., Karolyi G.A., Longstaff F.A., Sanders A.B. [1992] « An empirical comparison of alternative models of short-term interest rate », *Journal of Finance*, 47:1209-1227.

Chourdakis K. [2004] « Characteristic functions and Regime Switching Models », *CCFEA Workshop on Computational Finance* 2004, http://thePonyTail.net

Cohen G., Culioli J.-C. [1994] *Optimisation stochastique sous contraintes en espérance*, Rapport interne Centre Automatique et Systèmes, Ecole des Mines de Paris, no. A–288.

Consigli G., Dempster M.A.H. [1998] *The CALM stochastic programming model for dynamic asset-liability management*, dans Worldwide Asset and Liability Modelling, Ziemba & Mulvey, Cambridge University Press, Cambridge , UK, pp 464-500.

COSP - Stochastic Programming Community : http://stoprog.org/ [17, 101]

Cox J.C., Huang C.F. [1989] « Optimum Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process », *Journal of Economic Theory*, 49, 33-83.

Cox J.C., Ingersoll J.E., Ross S.A. [1985] « A Theory of the Term Structure of Interest Rates », *Econometrica*, 53:385-407.

Dantzig G.B. [1955] « Linear programming under uncertainty », *Management Science* 1, (1955) pp. 197-206.

Dantzig G.B., Glynn P.W. [1990] « Parallel processors for planning under uncertainty », *Annals of Operations Research*, 22, 1-22.

Date P., Wang C. [2009] « Linear Gaussian Affine Term Structure Models with Unobservable Factors: Calibration and Yield Forecasting », *European Journal of Operations Research*, 195 : 156-166 BURA.

Daykin C.D., Hey G.B. [1990] « Managing uncertainty in a general insurance company », *Journal of the Institute of Actuaries*, 117, 173-277

De Berg M., Cheong O., van Kreveld M., Overmars M. [2008] *Computational Geometry : Algorithms and Applications*, Springer-Verlag.

Décamps J.P. [1993] « Une Formule Variationnelle pour les Obligations du Secteur Privé », *Finance*, 14, 2, 61-77.

Delarue A. [2001] « Evaluation des réserves prudentielles en répartition », Lettre de l'Observatoire des retraites, n°12 mars 2001.

Deler B., Nelson B.L. [2001] « Modeling and generating multivariate time series with arbitrary marginals and autocorrelation structures », *Proceedings of the 2001 Winter Simulation Conference*, Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, 275-282.

Dempster M., Evstigneev I., Schenk-Hoppé K. [2003] « Exponential Growth of Fixed Mix Assets in Stationary Markets », *Finance and Stochastics*, 7, 263-276.

Dert C. [1995] *A dynamic model for asset/ liability management for defined benefit pension funds*, dans Worldwide Asset and Liability Modelling, Ziemba & Mulvey, Cambridge University Press, Cambridge , UK, pp 501-536.

Dogan K., Goetschalkx M. [1999] « A primal decomposition method for the integrated design of multi-period production–distribution systems », *IIE Transactions*, Volume 31, Number 11, November 1999 , pp. 1027-1036(10).

Dupacova J., Consigli G., Wallace S. [2000] « Generating scenarios for multistage stochastic programs », *Annals of Operations Research*, 100 (2000) 25-53 [17]

Dupacova J., Polivka J. [2009] « Asset-liability management for Czech pension funds using stochastic programming », *Annals of Operations Research*. Volume 165, Number 1 / janvier 2009, pages 5-28.

Durbin J., Watson G. S. [1951] « Testing for Serial Correlation in Least Squares Regression, II » *Biometrika* 38, 159–179.

Dyson A.C.L., Exley C.G. [1995] « Pension fund asset valuation and investment », *British Actuarial Journal*, 1: 471-557.

Elton E. J., Gruber M.J. [1973] « Estimating the dependence structure of share prices – implications for portfolio selection », *Journal of Finance*, 28, 1203-1232.

Emmerich M.T.M. [2005] *Single and Multi-objective evolutionary design optimization assisted by Gaussian Random Field Metamodels,* Thèse de doctorat, Université de Dortmund, Dortmund.

Engle R.F. [1982] « Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation », *Econometrica*, 50, 987-1007.

Escudero L.F., Kamesam P.V., King A., Wets R.J.B. [1993] « Production planning via scenario modeling », *Annals of Operations Research*, 43, 311-335.

Faleh A. [2011] «Un modèle de programmation stochastique pour l'allocation stratégique d'actifs d'un régime de retraite partiellement provisionné », *soumis*.

Faleh A., Planchet F., Rullière D. [2010] « Les générateurs de scénarios économiques : de la conception à la mesure de la qualité », Assurances et gestion des risques, n° double avril/juillet, Vol. 78 (1/2).

Fama E.F., [1981] « Stock Returns, Real Activity, Inflation, and Money », *American Economic Review*, September, pp. 545-565.

Fisher I. [1930] *The Theory of Interest*, New York: Macmillan.

Fisher R.A. [1954] *Statistical Methods for Research Workers*, 12ème edition, Edinburgh, Oliver & Boyd, 339 p.

Föllmer H., Kabanov Y. [1998] « Optional Decomposition and Lagrange Multipliers ». *Finance and Stochastics*, Vol 2 No 1, 1998.

Fourer R., Gay D.M., Kernighan B.W. [2002] *AMPL : A Modelling Language for Mathematical Programming*, 2nd edition, Duxbury Press.

Frauendorfer K., Jacoby U., Schwendener A. [2007] « Regime switching based portfolio selection for pension funds », *Journal of Banking & Finance* 31 (2007) 2265–2280.

Friggit J. [2001] *Prix des logements, produits financiers immobiliers et gestion des risques*, Paris : Économica.

Friggit J. [2007] « Long Term (1800-2005) Investment in Gold, Bonds, Stocks and Housing in France – with Insights into the USA and the UK : a Few Regularities », *CGPC Working paper*. http://www.adef.org/statistiques/index.htm

Gibson M.S., Boyer B.H. [1998] « Evaluating forecasts of correlation using option pricing », *Journal of Derivatives*, 6, 13-38.

Ginsbourger D. [2009] *Multiples métamodèles pour l'approximation et l'optimisation de fonctions numériques multivariables*, Thèse de doctorat de mathématiques appliquées, Ecole nationale supérieure des mines de Saint-Etienne, no 519MA.

Godfrey L. [1979] « Testing the Adequacy of a Time Series Model », *Biometrika*, 64, 67-72.

Gordon M., Shapiro E. [1956] « Capital Equipment Analysis : The Required Rate of Profit », *Management Science*, 3, p. 102-110.

Hainaut D., Devolder P. [2005] « Management of a pension fund under a VaR constraint », *working paper*.

Hainaut D., Devolder P. [2007] « Management of a pension fund under mortality and financial risks », *Insurance: Mathematics and Economics*, 41 (2007) 134–155.

Hamilton J. D. [1989] « A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle », *Econometrica*, 57,357-384

Hamilton J. D., Susmel R. [1994] « Autoregressive Conditional Heteroskedasticity and Changes in Regime », *Journal of Econometrics*, 64, 307-333.

Hansen E.R. [1979] « Global optimization using interval analysis : the one dimensional case », *Journal of Optimization Theory and Applications*, 29 :331-344.

Hardy M.R., [2001] « A Regime Switching Model of Long-Term Stock Returns », *North American Actuarial Journal*, 5 (2), 41-53. http://www.soa.org/library/journals/north-american-actuarial-journal/2001/april/naaj0104_4.pdf

Heath D., Jarrow R.A., Merton A. [1992] « Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation », *Econometrica*, 60:77-105.

Hibbert J., Mowbray P., Turnbull C. [2001] *A Stochastic Asset Model & Calibration for Long-Term Financial Planning Purposes*, Rapport Barrie & Hibbert Limited. http://www.actuaries.org.uk/__data/assets/pdf_file/0014/26312/hibbert.pdf

Hicks J.R. [1946] *Value and capital*, Second Edition, Oxford : Clarendon Pressa. Edition française publiée en 1956, "Valeur et Capital", Dunod.

Hilli P., Koivu M., Pennanen T., Ranne A. [2007] « A stochastic programming model for asset and liability management of a Finnish pension company », *Annals of Operations Research*, 152(2007), pp. 115-139

Ho T.S.Y., Lee S.B. [1986] « Term Structure Movements and Pricing Interest Rate Contingent Claims », *Journal of Finance*, 41:1011-1029.

Hochreiter R., Pflug G. Ch. [2002] *Scenario tree generation as a multidimensional facility location problem*, Rapport technique *AURORA*, TR2002-33, December 2002.

Holmer M. [1995] *Integrated asset/ liability management: an implementation case study*, dans Worldwide Asset and Liability Modelling, Ziemba & Mulvey, Cambridge University Press, Cambridge , UK.

Horst R., Parlados P.M. [1995] « Handbook of Global Optimization », Kluwer Academic Publishers, Dordrecht Boston London.

Horta P., Mendes C., Vieira I. [2008] « Contagion effects of the US Subprime Crisis on Developed Countries », CEFAGE-UE, *Working Paper* 2008/08.

Hoyland K., Kaut M., Wallace S.W. [2003] « Heuristic for moment-matching scenario generation », *Computational Optimization and Applications*, 24(2-3):169–185, 2003.

Huber P. [1995] « A review of Wilkie's stochastic asset model », *British Actuarial Journal*, 1, 181-211.

Hull J. [2000] *Options, Futures, and Other Derivative Securities*, New Jersey: Prentice-Hall, 5th. Edition.

Hull J.C., White A. [1990] « Pricing Interest Rate Derivative Securities », *Review of Financial Studies*, 3:573-592.

Hull J.C., White A. [1994] « Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models », *Journal of Derivatives* (Winter), 37-48.

Jarque C.M., Bera A.K. [1980] « Efficient tests for normality, homoscedasticity and serial independence of regression residuals », *Economics Letters* 6 (3): 255–259.

Johnson N. L. [1949] « Systems of frequency curves generated by methods of translation », *Biometrika* 36 (1): 297-304.

Jones D.R., Pertunen C.D., Stuckman B.E. [1993] « Lipschitzian optimization without the Lipschitz constant », *Journal of Optimization Theory and Applications*, 79[1], 157-181.

Jones D.R., Schonlau M., Welch W.J. [1998] « Efficient global optimization of expensive blackbox functions », *Journal of Global Optimization*, 13, 455-492.

Kallberg J.G., White R.W., Ziemba W.T. [1982] « Short Term Financial Planning under Uncertainty », *Management Science*, Vol. 28, No. 6, June 1982, pp. 670-682.

Kaut M., Wallace S.W., [2003] « Evaluation of scenario-generation methods for stochastic programming », SPEPS, *Working Paper*, 14 (http://edoc.hu-berlin.de/series/speps/2003-14/PDF/14.pdf), 2003.

Kharoubi-Rakotomalala C. [2008] *Les fonctions copules en finance*, Publications De La Sorbonne, collection Sorbonensia Oeconomica.

Kiefer J., Wolfowitz J. [1952] « Stochastic estimation of the maximum of a regression function », *Annals of Mathematical Statistics*, 23, 462-466.

Kim D., Santomero A. [1988] « Risk in banking and capital regulation », *Journal of Finance*, Vol. 43 (5), p.1219-1233.

Kim J.R. [2003] « The stock return-inflation puzzle and asymmetric causality in stock returns, inflation and real activity », *Discussion paper*, Economic Research Centre of the Deutsche Bundesbank, 03/03, January.

Klassen P. [1997] « Solving stochastic programming models for asset/ liability management using iterative disaggregation », *Research Memorandum, 0011,* VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics.

Kouwenberg R. [2001] « Scenario Generation and Stochastic Programming Models for Asset Liability Management », *European Journal of Operational Research*, vol. 134, 51-64.

Krige D.G. [1951] *A statistical approach to some mine valuations and allied problems at the Witwatersrand*, Thèse de master de l'Université de Witwatersrand.

Kusy M.I., Ziemba W.T. [1986] « A Bank Asset and Liability Management Model », *Operations research*, Vol. 34, No. 3, May-June 1986, pp. 356-376.

Lawler E.L., Wood D.E. [1966] « Branch and Bound methods : a survey », *Operations Research*, Vol. 14, no 4, pp 699-719.

Leibowitz M.L., Kogelman S., Bader L.N. [1992] « Asset Performance and Surplus Control: A Dual-Shortfall Approach », *The Journal of Portfolio Management,* WINTER.

Le Vallois F., Palsky P., Paris B., Tosetti A. [2003] *Gestion actif passif en assurance vie : Réglementation - Outils – Méthodes*, Economica.

Macaulay F.R. [1938] « The Movements of Interest Rates. Bond Yields and Stock Prices in the United States since 1856 », New York: *National Bureau of Economic Research*.

Mandelbrot B. [2005] *Une approche fractale des marchés : risquer, perdre et gagner*, Editions Odile Jacob.

Malliaris A. G., Brock W.A. [1982] *Stochastic Methods in Economics and Finance*, New York: Elsevier Science Publishers.

Markowitz H. M. [1952] « Portfolio Selection », *The Journal of Finance*, Vol. 7, No.1, March.

Martellini L. [2006] *Managing Pension Assets: from Surplus Optimization to Liability-Driven Investment,* EDHEC Risk and Asset Management Research Centre.

Martellini L., Priaulet P., Priaulet S. [2003] *Fixed-Income Securities: Valuation, Risk Management and Portfolio Strategies*, Wiley, 2003.

Mathias K., Whitley D., Kusuma A., Stork C. [1996] « An empirical evaluation of genetic algorithms on noisy objective functions », *Genetic Algorithms for Pattern Recognition* S.K. Pal, ed. pp: 65-86. CRC Press, 1996.

Mazzola J.B., Schantz R.H. [1996] « Multiple-facility loading under capacity-based economies of scope », *Naval Research Logistics* (NRL), Volume 44 Issue 3, Pages 229 – 256.

Merton R. [1971] « Optimum Consumption and Portfolio Rules in a Continuous Time Model », *Journal of Economic Theory*, 3, 373-413.

Merton R. [1976] « Option pricing when underlying stock returns are discontinuous », *Journal of Financial Economics*, 3, pp.125-144.

Merton R. [1990] *Continuous-Time Finance*, Oxford, UK.

Meucci A. [2005] *Risk and asset allocation*, New York : Springer.

Mitra S. [2006] « A White Paper on Scenario Generation for Stochastic Programming », *OptiRisk Systems:White Paper Series*, juillet.

Mulvey J..M., Thorlacius A.E. [1998] *The Towers Perrin Global Capital Market Scenario Generation System: CAP:Link*, dans Worldwide Asset and Liability Modelling, Ziemba & Mulvey, Cambridge University Press, Cambridge , UK.

Munk C., Sorensen C., Vinther T.W. [2004] « Dynamic asset allocation under mean-reverting returns, stochastic interest rates and inflation uncertainty », *International Review of Economics & Finance*, Volume 13, Février 2004,141-166.

Nelder J., Mead R. [1965] « A simplex method for function minimization », *Computer Journal*, vol. 7, no 4, p.308-313.

Norkin V., Pflug G.Ch., Ruszczynski A. [1996] « A branch and bound method for stochastic global optimization », *Mathematical Programming*, vol 83, no 1-3, pp 452-450.

Pennanen T., Koivu M. [2002] « Integration quadratures in discretization of stochastic programs », *Stochastic Programming*, E-Print Series, http://www.speps.info , May 2002.

Perold A. F, Sharpe W. F. [1988] « Dynamic Strategies for Asset Allocation », *Financial Analysts Journal*, 44, no. 1.

Pierre S.N. [2009] *Passif social, construction du portefeuille d'investissement et couverture du risque de taux* , Mémoire d'actuaire, CNAM.

Pirkul H., Jayaraman V. [1996] « Production, Transportation, and Distribution Planning in a Multi-Commodity Tri-Echelon System », *Transportation Science*, *30*(4), 291-302.

Planchet F. [2009] « Quel modèle d'actifs en assurance ? », *la Tribune de l'Assurance* (rubrique « le mot de l'actuaire »), n°136 du 01/05/2009.

Planchet F., Thérond P.E. [2005] *Modèles financiers en assurance – Analyses de risque dynamiques, Economica*.

Planchet F., Thérond P.E. [2007] *Pilotage technique d'un régime de rentes viagères, Economica*.

Planchet F., Thérond P.E., Kamega A. [2009] *Scénarios économiques en assurance - Modélisation et simulation*, Paris : Economica.

Portrait R., Poncet P. [2008] *Finance de marché – Instruments de base, produits dérivés, portefeuilles et risques*, Dalloz.

Rambaruth G. [2003] *A comparison of Wilkie-type stochastic investment models,* Mémoire de master, City University,UK.

Ranne A. [1998] « The Finnish Stochastic Investment Model », *Transactions of 26th ICA*, 213-238.

Rasmussen K.M., Clausen J. [2004] « Mortgage Loan Portfolio Optimization Using Multi Stage Stochastic Programming », *Journal of Economic Dynamics and Control*, 31, 742-766.

Redington F.M. [1952] « Review of the principles of life office valuations », *Journal of the Institute of Actuaries*, 78: 1-40.

Robbins H., Monro S. [1951] « A Stochastic approximation method », *Annals of Mathematical Statistics*, 22, 400-407.

Roncalli T. [1998] *La structure par terme des taux zéro : modélisation et implémentation numérique*, Thèse de doctorat, Université de Bordeaux IV.

Roy A.D. [1952] « Safety-first and the holding of asset », *Econometrica*, 20, 431-449.

Rudolf M., Ziemba W.T. [2004] « Intertemporal surplus management », *Journal of Economic Dynamics & Control*, 28 (2004) 975 – 990.

Rullière D., Faleh A., Planchet F. [2010] « Un algorithme d'optimisation par exploration sélective », *soumis*.

Rullière D., Faleh, A., Planchet F. [2010] « Fonction d'information et optimisation globale d'une fonction bruitée », *preprint*.

Santner T., Williams B., Notz W. [2003] *The design and analysis of computer experiments*, Springer Verlag, New York.

Santosa F., Alexandrov O. [2005] « A topology-preserving level set method for shape optimization », *Journal of Computational Physics*, Volume 204 , Issue 1 (March 2005), Pages: 121 – 130.

Schubert B. [1972] « A sequential method seeking the global maximum of a function », *SIAM Journal on Numerical Analysis*, 9 :379-388.

Sharpe W.F., Tint L.G. [1990] « Liabilities - A new approach », *Journal of Portfolio Management,* Winter, 1990, 5-10.

Shapiro A., Verweij B., Shabbir A., Kleywegt A.J., Nemhauser G. [2003] « The Sample Average Approximation Method Applied to Stochastic Routing Problems: A Computational Study », *Computational Optimization and Applications*, Volume 24, Numbers 2-3 / février 2003,pages 289-333.

Shiller R.J. [1981] « Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? », *NBER Working Papers* 0456, National Bureau of Economic Research, Inc.

Skintzi V.D., Spyros X.S. [2007] « Evaluation of correlation forecasting models for risk management », *Journal of Forecasting*, 26:7, 497-526.

Smith N.A., Tromble R.W. [2004] *Sampling uniformly from the unit simplex*, Rapport technique, Johns Hopkins University.

Smith A.D. [1996] « How actuaries can use financial economics », *British Actuarial Journal*, 2: 1057-1193.

Sommerville D.M.Y. [1958] *An introduction to the Geometry of n dimensions*, New York, Dover edition.

Strugarek, C. [2006] *Approches variationnelles et autres contributions en optimisation stochastique*, ENPC, Thèse de doctorat.

Student (ou Gosset W.S.) [1908] « The Probable Error of a Mean », *Biometrika*, 6, 1–25.

Tanaka S., Inui K. [1995] « Modelling Japanese financial markets for pension ALM simulations », *5th AFIR colloquium*, 563-584

Thomson R.J. [1994] « A stochastic investment model for actuarial use in South Africa », *Transactions of the Actuarial Society of South Africa 1*.

Tuckman B., [2002] *Fixed Income Securities: Tools for Today's Markets*, John Wiley & Sons, Inc, 10-01-2002, 528 pages.

Vasicek O. [1977] « An equilibrium Characterization of the Term structure », *Journal of Financial Economics*, 5:177-188.

Villemonteix J. [2009] *Optimisation de fonctions coûteuses*, Thèse de doctorat de physique, Université Paris Sud 11, Faculté des sciences d'Orsay, no 9278.

Wainscott C. B. [1990] « The stock-bond correlation and its implications for asset allocation », *Financial Analysts Journal*, 46, 55-60.

Waring B. [2004] « Liability-Relative Investing II: Surplus Optimization with Beta, Alpha, and an Economic View of the Liability », *working paper*, September 2004.

Wilkie D. [1984] « Steps towards a comprehensive stochastic model », *Occasional Actuarial Research Discussion Paper*, The Institute of Actuaries, London, 36:1-231.

Wilkie D. [1986] « A Stochastic Investment Model for Actuariel Use », *Transactions of the Faculty of Actuaries*, 39:341-403.

Wilkie D. [1995] « More on a Stochastic Model for Actuarial Use », *British Actuarial Journal*, pp. 777-964. http://www.ressources-actuarielles.net/jwa/documentation/1226.nsf.

Wolfe M.A. [1996] « Interval Methods for Global Optimization », *Applied Mathematics and Computation*, Vol. 75, Issues 2-3, pp. 179-206.

Yen S.H., Hsu Ku Y.H. [2003] « Dynamic Asset Allocation Strategy for Intertemporal, Pension Fund Management with Time-Varying Volatility », *Academia Economic Papers*, 31: 3 (September 2003), 229–261.

Zadeh L.A. [1965] « Fuzzy Set », *Information and Control*, 8 [3] : 338-353.

Zenios S.A. [2007] *Practical Financial Optimization : Decision Making For Financial Engineers*, Blackwell Publishers, Oxford, 2007.

Ziemba W.T. [2003] *The Stochastic Programming Approach to Asset-Liability and Wealth Management*, AIMR-Blackwell.

Ziemba W. T., Mulvey J.M. [1998] *Worldwide Asset and Liability Modelling*, Cambridge University Press, UK.

Table des matières

Sommaire…….....1

Introduction générale……4

Partie I : Les générateurs de scénarios économiques (GSE)……...…..17

Introduction ……..…….…….....18

Chapitre 1 : Etude des GSE…….……..19

I- Présentation théorique ….…….. ……..……19

I-1 Problèmes théoriques de modélisation liés aux GSE….……..……...19

I-1-1 Modèles des taux d'intérêt ……..…….19

I-1-1-1 *Les modèles d'absence d'opportunité d'arbitrage (AOA)...……..19*
I-1-1-2 *Les modèles d'équilibre général…….…….…….21*

I-1-2 Modèles de rendement des actions ……..……....23
I-1-3 Autres classes d'actifs …….……...24

I-2 Littérature sur les structures des GSE…….……..…….……24

I-2-1 Modèles à structure par cascade……..…….……25
I-2-2 Modèles basés sur les corrélations…….……..27

II- Mise en œuvre d'un GSE…….……...……...30

II-1 Structure schématique de projection de scénarios pour un GSE…….……..…….30
II-2 Méthodologies pour la génération de scénarios économiques……..…….…….34

II-2-1 Les approches basées sur l'échantillonnage…….……..…….....34
II-2-2 Les approches basées sur le *matching* des propriétés statistiques…….…….....35
II-2-3 Les approches basées sur les techniques de *Bootstrapping*….…….……..36
II-2-4 L'utilisation de l'Analyse en Composantes Principales…….……..……...36

II-3 Probabilité risque-neutre Vs probabilité réelle…….……..36
II-4 L'exclusion de valeurs négatives pour les modèles de taux…….……..37

III- Mesure de qualité…….……..……..38

III-1 Mesure qualitative de la qualité d'un GSE……38
III-2 Mesure quantitative de la qualité d'un GSE……..40

III-2-1 Présentation du problème……..40
III-2-2 Le test de la qualité des décisions obtenues par le GSE…….……...42

III-2-2-1 *La condition de stabilité* …….……..*42*
III-2-2-2 *La condition d'absence de biais*…….……...*44*

III-3 Application numérique...…….…….……....45

Conclusion du chapitre 1……...49

Chapitre 2 : Construction d'un générateur de scénarios économiques…….……50

I- Conception et composantes ……..…….…..……50

I-1 Conception théorique ……..……..…….……...50

I-2 Composantes ……...…….……...53

I-2-1 Structure de corrélation …….……..53
I-2-2 Choix des modèles de diffusions…….……58

I-2-2-1 *Le modèle de structure par terme des taux d'intérêt…….…….. …..58*
I-2-2-2 *Le modèle des actions…….……..… .65*
I-2-2-3 *Le modèle des obligations "crédit"……..…….……..….. 67*
I-2-2-4 *Le modèle de l'immobilier…….....71*
I-2-2-4 *Le monétaire…….…..... 71*

II- Calibrage…… ……..……73

II-1 Présentation du contexte de l'étude ……..……..……...73

II-1-1 Description des données retenues…….…..……..….73
II-1-2 Approche et méthode de calibrage…….…..…….…75

II-2 Calibrage du modèle de l'inflation…...…….……..……...76

II-2-1 Méthode…….……..76
II-2-2 Données…….……..77
II-2-3 Estimation des paramètres…….…….…….....77
II-2-4 Tests d'adéquation du modèle…….……....78

II-3 Calibrage du modèle des taux réels...……..……..…….79

II-3-1 Méthode…….…….……..……..79
II-3-2 Données…….…….……..... .80

II-3-3 Estimation des paramètres……..……..……...80
II-3-4 Tests d'adéquation du modèle…..……...……..81

II-4 Calibrage du modèle des actions..……..……83

II-4-1 Méthode……..……...……..83
II-4-2 Données……..……..…….…….....88
II-4-3 Estimation des paramètres…..…….…….....88

II-5 Calibrage du modèle de l'immobilier……..…….…….....89

II-5-1 Méthode..……..…….....89
II-5-2 Données…….……..…...89
II-5-3 Estimation des paramètres…….……..…….....89
II-5-4 Tests d'adéquation du modèle……..….....90

II-6 Calibrage du modèle de crédit...……...91

III- Projection des scénarios ……..……93

III-1 Éléments sur la mise en œuvre……....…….……..93
III-2 Tests du modèle : paramètres et matrice de corrélation…….……....96

III-2-1 Tests sur les paramètres et sur les développements informatiques…….…….....97
III-2-2 Tests sur les matrices de corrélation……...….....…….98

III-3 Résultats de projection……..……..……..……....99

Conclusion du chapitre 2…….110

Conclusion de la partie I……...110

Partie II : L'allocation stratégique d'actifs dans le cadre de la gestion actif-passif (ALM)……..……..112

Introduction……...113

Chapitre 1 : Les modèles d'ALM classiques (déterministes)……118

I- Immunisation du portefeuille…..……118

I-1 Adossement des flux de trésorerie……..……...119
I-2 Adossement par la duration……..…….120

II- Modèles basés sur le surplus…….…….125

II-1 Modèle de Kim et Santomero [1988]...……..……..125
II-2 Modèle de Sharpe et Tint [1990] …….……...128
II-3 Modèle de Leibowitz [1992]…….……...130
II-4 Duration du Surplus…….……137

Conclusion du chapitre 1…….139

Chapitre 2 : Allocation stratégique d'actifs et stratégie *Fixed-Mix*…….....……140

I- Etude de l'allocation d'actifs dans le cadre de la stratégie *Fixed-Mix*…...……..143

I-1 Présentation…..……. ..143
I-2 Application…...……143

I-2-1 Modélisation du régime-type…...……...144

I-2-1-1 *Notations*…...……..…….*144*
I-2-1-1 *Modélisation*…….……*144*

I-2-2 Etude des critères d'allocation…...…….....146

I-2-2-1 A*nalyse générale de la solvabilité du régime*…...……..*147*
I-2-2-*2 Critères d'allocation en fonction du type du régime* ….....….*153*
I-2-2-3 *Critères d'allocation en fonction da la phase dans laquelle évolue le régime*…..……..…...……..*153*

I-2-3 Etude de la question du taux d'actualisation des engagements…...…...153
I-2-4 Hypothèses du modèle…...……. …...155
I-2-5 Méthodologie retenue…...……..158
I-2-6 Résultats obtenus ….……..……161

II- Proposition de méthodes numériques ……...……....168

II-1 Introduction…….…….168

II-1-1 Le problème d'optimisation……..168
II-1-2 Approche retenue……. 172

II-2 Une grille à pas variable par subdivision systématique……...173

II-2-1 Forme des zones de recherche…….….173
II-2-2 Choix de la zone à explorer ou segmenter……....175
II-2-3 Choix des paramètres (σ_K , α) du potentiel……..181
II-2-4 Choix du sommet de scission……183
II-2-5 Schéma de l'algorithme de scission systématique……184
II-2-6 Résultat final et critère de convergence…….…...185

II-3 Une grille à pas variable avec retirage possible…….……..188

II-3-1 Critères de scission……..……..188
II-3-2 Schéma de l'algorithme à retirage possible……..190

II-4 Applications numériques…..…….191

II-4-1 Fonction test utilisée…….……191
II-4-2 Comportement de l'algorithme à scission systématique……...193
II-4-3 Comportement de l'algorithme avec retirages possibles…….….199
II-4-4 Estimation des paramètres (σ_K , α)…….……...……..200
II-4-5 Critères de convergence……....203
II-4-6 Comparaison avec l'algorithme de Kiefer-Wolfowitz-Blum…..……...208

II-5 Conclusion……...……....211

Conclusion du chapitre 2…….……212

Chapitre 3 : Allocation stratégique d'actifs et modèles d'ALM dynamique……..…….213

I - L'allocation stratégique d'actifs dans le cadre des modèles classiques d'ALM dynamique…….……214

I-1 Présentation générale……...……. 214
I-2 Techniques d'assurance de portefeuille …….……...215

I-2-1 Principe…...……..…….215
I-2-2 Illustration…..……..……..216

I-2-2-1 *Contexte*……..……..…….....*216*
I-2-2-2 *Modélisation de l'actif*……..……..…….…...*217*
I-2-2-3 *Modélisation du passif*…..……..…….……..*217*
I-2-2-4 *Calibrage des paramètres de marché*…..……..……*218*
I-2-2-5 *Calibrage des paramètres de gestion*…..……..…..*218*
I-2-2-6 *Résultats*……..……*219*

I-3 Programmation dynamique……...……...……...…..223

II- L'allocation stratégique d'actifs dans le cadre de la programmation stochastique …….……....227

II-1 Introduction ……...……..……...……..227
II-2 Présentation de la programmation stochastique...……...…….228
II-3 Illustration de la programmation stochastique avec recours dans le cas de la planification de la production…….……..230

II-3-1 Cas mono-périodique (à deux étapes) ……..…..……..230
II-3-1 Cas multi-périodique (multi- étapes) ……..…..……232

II-4 Formulation mathématique de la programmation stochastique avec recours…...236
II-5 Résolution des programmes stochastiques…….……...……240

III- Nouvelle approche d'ALM par la discrétisation des scénarios économiques……..…….……....242

III-1 Méthode des quantiles de référence pour le GSE..……...……..242
III-2 Modèle d'optimisation basé sur la programmation stochastique.....…...……247
III-3 Discussion des résultats……. …….……...……253

Conclusion du chapitre 3…….261

Conclusion de la partie II ……261

Conclusion générale…….262

Index des graphiques……...…268

Bibliographie……....272

Table des matières……...284