



**HAL**  
open science

# Méthodes de Monte Carlo multi-niveaux pour l'estimation de paramètres statistiques de champs discrétisés en géosciences

Jérémy Briant

► **To cite this version:**

Jérémy Briant. Méthodes de Monte Carlo multi-niveaux pour l'estimation de paramètres statistiques de champs discrétisés en géosciences. Mathématiques générales [math.GM]. Université de Toulouse, 2024. Français. NNT : 2024TLSEP105 . tel-04868031

**HAL Id: tel-04868031**

**<https://theses.hal.science/tel-04868031v1>**

Submitted on 6 Jan 2025

**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.

# Doctorat de l'Université de Toulouse

préparé à Toulouse INP

---

Méthodes de Monte Carlo multi-niveaux pour l'estimation de  
paramètres statistiques de champs discrétisés en géosciences

---

Thèse présentée et soutenue, le 12 novembre 2024 par

**Jérémy BRIANT**

## École doctorale

EDMITT - École Doctorale Mathématiques, Informatique et Télécommunications de Toulouse

## Spécialité

Mathématiques et Applications

## Unité de recherche

IRIT - Institut de Recherche en Informatique de Toulouse

## Thèse dirigée par

Serge GRATTON et Anthony WEAVER

## Composition du jury

M. Olivier LE MAÎTRE, Président, CNRS Paris-Centre

M. Laurent DEBREU, Rapporteur, Inria Grenoble - Rhône-Alpes

Mme Clémentine PRIEUR, Examinatrice, Université Grenoble Alpes

M. Reda EL AMRI, Examineur, IFP Energie Nouvelles

Mme Selime GÜROL, Examinatrice, CERFACS

M. Serge GRATTON, Directeur de thèse, Toulouse INP

M. Anthony WEAVER, Directeur de thèse, CERFACS

## Membres invités

M. Paul Mycek, CERFACS

M. Ehouarn Simon, Toulouse INP



# Résumé

Les méthodes de réduction de variance multi-fidélités pour l'estimation statistique sont utilisées dans de plus en plus de domaines comme une amélioration par rapport à la méthode Monte Carlo. Dans leur forme originale, elles sont conçues pour estimer des paramètres statistiques scalaires. Cependant, dans certaines applications, par exemple en géosciences, les quantités d'intérêt dont on souhaite estimer les paramètres statistiques peuvent être des vecteurs ou des champs aléatoires. Dans le cas d'un vecteur aléatoire représentant un champ aléatoire discrétisé, l'utilisation d'un estimateur multi-fidélité n'est pas triviale, notamment si les entrées et sorties des différents modèles n'ont pas les mêmes dimensions sur les différents niveaux de fidélité. Des opérateurs de transfert entre les différents niveaux doivent alors être introduits. Cette thèse porte sur l'adaptation de l'estimateur Monte Carlo multi-niveaux (MLMC) à l'estimation de champs discrétisés en géosciences et sur l'analyse des opérateurs de transfert, permettant le choix d'opérateurs qui réduisent la variance du MLMC.

Afin de mieux comprendre l'effet des opérateurs de transfert, une analyse spectrale de l'estimateur MLMC est proposée. Des expériences numériques sur des problèmes simplifiés ont mis en évidence que le MLMC détériore l'estimation des composantes hautes fréquences d'un champ discrétisé comparé à un simple estimateur Monte Carlo. Une analyse théorique portant sur une classe de problèmes spécifique, similaire aux analyses développées pour les méthodes multigrilles, a permis une meilleure compréhension des disparités dans l'estimation, par MLMC, des différentes composantes d'un champ discrétisé. Ces différents résultats nous ont conduit à modifier les opérateurs de transfert utilisés afin d'y ajouter une étape de filtrage inspirée des méthodes multigrilles. L'ajout de filtres permet d'améliorer l'estimation des composantes hautes et basses fréquences, réduisant ainsi la variance totale de l'estimateur.

L'utilisation de méthodes de type multilevel best linear unbiased estimator (MLBLUE) a permis de mieux choisir les opérateurs de transfert et les filtres en vue de réduire la variance de l'estimateur. Ces différentes améliorations ont été appliquées à l'estimation de la variance d'un champ discrétisé obtenu par application d'un opérateur de covariance basé sur la diffusion. Les expériences numériques ont pu confirmer les résultats théoriques sur un problème plus complexe et mettre en évidence les améliorations apportées par rapport à un simple MLMC.

Cette recherche contribue à élargir les applications du MLMC à d'autres domaines, notamment les géosciences, et offre une meilleure compréhension de la méthode grâce aux liens tissés avec les méthodes multigrilles. Elle permet également d'améliorer l'estimateur MLMC en soulignant l'importance du choix des opérateurs de transfert.



# Remerciements

Au cours de ces quatre années de thèse, j'ai eu le plaisir d'être encadré par une équipe aussi grande que bienveillante, et je souhaite les remercier amplement. Tout d'abord, je tiens à remercier mon directeur de thèse, Serge Gratton, qui, grâce à sa vision globale du domaine et ses réflexions toujours pertinentes, a énormément contribué à l'élaboration de ce sujet de thèse et m'a aidé à bien démarrer ces années de recherche. Je remercie également Anthony Weaver, mon codirecteur de thèse, dont la bonne humeur au quotidien apporte toujours le sourire. Il a su faire en sorte que le sujet reste toujours aligné avec la pratique, donnant ainsi bien plus de sens à ce que j'ai pu produire. Je remercie tout particulièrement Paul Mycek qui a été mon principal encadrant côté Cerfacs et qui m'a apporté un soutien quasi quotidien durant ces quatre années. Il m'a enseigné de nombreuses qualités nécessaires en recherche, et m'a également transmis son intérêt profond pour le sujet, ce qui a été crucial pour garder la motivation sur une aussi longue période. Côté IRIT, c'est Ehouarn Simon qui m'a encadré et je le remercie chaleureusement. Ses conseils et retours ont toujours été les bienvenus et m'ont permis, d'une part d'atteindre une vision plus large du sujet, et d'autre part de me sentir plus serein face à ce qu'il restait à faire. Enfin, je remercie Selime Guröl dont tous les retours m'ont grandement aidé à mieux expliquer et présenter mes travaux et dont les encouragements réguliers m'ont fait chaud au cœur. J'ai été chanceux d'être entouré par toutes ces personnes qui m'ont toutes aidées, à un moment ou un autre, à reprendre confiance en mes capacités à terminer cette thèse avec succès.

J'aimerais également remercier Olivier Le Maître et Laurent Debreu qui ont accepté d'être rapporteurs sur ce manuscrit. Le temps qu'ils ont consacré à l'évaluer m'a apporté un retour extérieur sur mes travaux après quatre ans de thèse, ce qui est très précieux. Je remercie particulièrement Olivier d'être venu à Toulouse pour présider la soutenance. J'aimerais aussi remercier les examinateurs : Clémentine Prieur et Reda El Amri pour les différentes questions et discussions. Ce moment d'échange fut enrichissant et m'a permis de prendre bien plus de recul sur la recherche effectuée durant cette thèse.

En travaillant dans deux laboratoires différents j'ai aussi eu la chance de découvrir deux équipes de travail aussi accueillantes et chaleureuses l'une que l'autre. Ce fut un plaisir de côtoyer toute l'équipe Algo du Cerfacs, les permanents évidemment mais aussi tous ceux qui n'y ont fait qu'un passage le temps d'un stage, d'une thèse ou d'un post-doc. Je remercie particulièrement Mayeul et Olivier avec qui j'ai eu le plaisir de travailler, mais aussi de faire des sudokus durant les pauses café. À l'IRIT, je remercie tout le troisième étage du bâtiment F, que ce soit l'équipe APO dont je faisais partie et où la bonne ambiance et la solidarité entre doctorants est une belle source de motivation, mais aussi l'équipe ACADIE avec qui j'ai pu partager de nombreuses discussions sur des sujets très, très variés.

Quand il s'agissait de penser à autre chose que la thèse, j'étais entouré de tout un groupe d'amis de l'ENSEEIHHT que je remercie pour tous les moments partagés et tous ceux à venir, autour de bières et de jeux de cartes. J'adresse une attention toute particulière à mon colocataire durant toute cette thèse, Yann, grâce à qui le quotidien

était bien plus vivant, ainsi qu'à Maxime qui nous faisait sortir afin de faire autre chose que des maths.

Finalement, je remercie toute ma famille qui arrive à me faire parvenir leur soutien, même quand il provient du bout de la Bretagne. Si j'ai réussi à atteindre le bout de toutes ces années d'études, c'est grâce à toutes celles qui ont précédé durant lesquelles ils m'ont tous aidé à progresser.



# Table des matières

<b>Résumé</b>	<b>ii</b>
<b>Remerciements</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Monte Carlo multi-niveaux</b>	<b>6</b>
2.1 Notations	6
2.2 Monte Carlo pour l'espérance d'une variable aléatoire réelle	6
2.3 MLMC pour l'espérance d'une variable aléatoire réelle	7
2.4 Illustration MC et MLMC	10
2.4.1 Présentation de l'illustration	10
2.4.2 Illustration MC	11
2.4.3 Illustration MLMC	12
2.5 Conclusion	13
<b>3 MLMC pour des champs discrétisés spatialement</b>	<b>15</b>
3.1 MLMC pour des champs discrétisés	15
3.1.1 Opérateurs de transfert	16
3.1.2 Formulation de l'estimateur MLMC pour des champs discrétisés	16
3.2 Formulation forte et discrétisation	17
3.2.1 Discrétisation de l'espace par grilles	18
3.2.2 Grilles centrées sur les cellules et opérateurs de transfert linéaires	18
3.3 Analyse spectrale	20
3.3.1 Base de Hartley	20
3.3.2 Effets des opérateurs de transfert sur la base de Hartley	23
Prolongation de la base de Hartley	24
Restriction de la base de Hartley	25
Combinaison de la restriction et de la prolongation	27
3.3.3 Impact sur la variance du MLMC	28
Modèles linéaires	28
Modèles linéaires, symétriques circulants	29
Opérateurs de Galerkin	30
3.4 Illustration 1D	33
3.4.1 Présentation du problème	33
3.4.2 Résultats	35
3.5 Conclusion	40
<b>4 Filtered-MLMC</b>	<b>42</b>
4.1 Ajout des filtres	42
4.1.1 Estimateur F-MLMC	43
4.1.2 Filtres de Shapiro	43
4.2 Analyse spectrale	44

4.2.1	Effets de l'opérateur de filtrage sur la base de Hartley . . . . .	44
4.2.2	Impact sur la variance du F-MLMC . . . . .	45
	Modèles linéaires . . . . .	45
	Modèles linéaires, symétriques circulants . . . . .	46
	Opérateurs de Galerkin . . . . .	47
4.3	Illustration 1D . . . . .	49
4.4	Choix du filtre et des opérateurs d'interpolation . . . . .	56
4.4.1	Augmentation de l'ordre des opérateurs de transfert . . . . .	56
4.4.2	Étude et comparaison de différents filtres . . . . .	57
4.4.3	Weighted MLMC . . . . .	60
	Liens entre F-MLMC et WMLMC . . . . .	60
	Illustration 1D WMLMC . . . . .	62
4.5	Conclusion . . . . .	66
<b>5</b>	<b>Expériences numériques</b>	<b>68</b>
5.1	Contexte et présentation du problème . . . . .	68
5.1.1	Problème de normalisation d'un opérateur de covariance basé sur la diffusion . . . . .	68
5.1.2	Construction de l'opérateur de diffusion . . . . .	69
5.1.3	Estimation des coefficients de normalisation par randomisation	71
5.2	Normalisation 1D . . . . .	71
5.2.1	Résultats du F-MLMC . . . . .	74
5.2.2	Résultats du WMLMC . . . . .	80
5.2.3	Conclusion . . . . .	83
5.3	Normalisation 2D . . . . .	84
5.3.1	Résultats du F-MLMC . . . . .	89
5.4	Conclusion . . . . .	99
<b>6</b>	<b>Conclusion et perspectives</b>	<b>101</b>
<b>A</b>	<b>Propriétés de la base de Hartley</b>	<b>105</b>
A.1	Orthogonalité de la matrice de Hartley . . . . .	105
A.2	Les matrices symétriques circulantes sont diagonalisables dans la base de Hartley . . . . .	105
	<b>Bibliographie</b>	<b>109</b>

## Chapitre 1

# Introduction

L'utilisation de simulateurs numériques pour modéliser un phénomène physique est omniprésente dans de nombreux domaines depuis déjà plusieurs décennies. La modélisation mathématique et numérique d'un système physique peut permettre de simuler son comportement afin de mieux l'étudier ou de prédire sa réponse à une situation précise. L'accès à une puissance de calcul de plus en plus importante a permis de faire évoluer ces simulateurs pour les rendre plus complets et précis. Cependant, des erreurs sur certains paramètres, certaines observations ou simplement la sensibilité du problème considéré peuvent aboutir à des comportements du modèle très éloignés de la réalité. Une manière de prendre en compte de telles erreurs est de modéliser les entrées par des variables aléatoires et de chercher, ensuite, à estimer des paramètres statistiques associés aux sorties du simulateur. La méthode Monte Carlo [9, 61, 44], de par sa simplicité et sa flexibilité, est l'une des plus populaire pour l'estimation statistique. Cette méthode se repose sur la génération d'un ensemble de réponses du simulateur considéré afin de pouvoir estimer les statistiques voulues. La complexité des simulateurs numériques peut alors devenir un inconvénient pour de telles méthodes, leur coût de calcul élevé empêchant la génération d'un ensemble trop grand et limitant donc fortement la précision des estimations. Un exemple important est celui des géosciences où la complexité des phénomènes physiques modélisés peut requérir de nombreuses heures de calculs pour une simple application du modèle. Il est ainsi possible en météorologie ou en océanographie de devoir se limiter à des ensembles de seulement quelques dizaines de membres pour estimer une quantité d'intérêt telle qu'une moyenne ou une matrice de covariance d'erreur, ce qui implique une importante erreur d'estimation.

Les méthodes multi-fidélités [57] cherchent à combiner le coût de calcul faible d'un simulateur peu complexe avec la précision d'un simulateur plus complet pour obtenir une estimation de meilleure qualité. Ces méthodes ont apporté de nombreuses améliorations dans différents domaines et notamment pour l'estimation statistique où elles peuvent permettre de considérer des ensembles de réponses plus grands en utilisant des modèles basse fidélité. Il existe de nombreuses manières de construire des simulateurs basse fidélité; les modèles de substitution (également appelés métamodèles, proxys, ou encore surfaces de réponse), par exemple, approchent les relations entre les entrées et sorties d'un modèle avec une représentation fonctionnelle. Les réseaux de neurones [68] en sont un bon exemple mais les polynômes du chaos [42, 47, 53, 77] ou les processus gaussiens [59, 60] sont aussi couramment utilisés. Une autre manière permettant de construire des simulateurs de basse fidélité est d'utiliser des discrétisations plus grossières de l'espace ou du temps. De tels modèles ont été fortement étudiés pour leurs utilisations dans les méthodes multigrilles [7, 69, 30] mais ils peuvent être utilisés dans de nombreux autres contextes. Ces simulateurs sont particulièrement bien adaptés à des applications en géosciences qui reposent souvent sur la résolution d'équations aux dérivées partielles par discrétisation de

l'espace ou du temps.

En fonction de la collection de modèles de différentes fidélités accessibles, il existe de nombreuses méthodes multi-fidélités pour l'estimation statistique. Les variables de contrôle (CV) [40, 41, 54, 43], par exemple, combinent le simulateur haute fidélité avec un ou plusieurs simulateurs basse fidélité dont l'espérance est connue. Les méthodes Monte Carlo multi-fidélité (MFMC) [55, 22] utilisent des estimations Monte Carlo [61, 44] sur des modèles de différentes fidélités afin de réduire l'erreur quadratique moyenne (mean square error ou MSE) d'un simple estimateur Monte Carlo. Parmi les méthodes multi-fidélités on distingue la sous-classe des méthodes multi-niveaux qui supposent l'accès à un ratio précision / coût pour chacun des modèles considérés. La connaissance d'un tel ratio permet de hiérarchiser précisément les différents modèles, ce qui n'est pas forcément le cas de toutes les méthodes multi-fidélités. La construction de modèles de basse fidélité par grilles de différentes résolutions est un exemple courant permettant d'appliquer ces méthodes multi-niveaux. On citera, par exemple, le filtre de Kalman d'ensemble multi-niveaux [34, 11] qui est utilisé dans les étapes d'assimilation de données [10], courantes en géosciences. Les méthodes Monte Carlo multi-niveaux [32, 24, 23] sont parmi les plus populaires depuis quelques années et utilisent des niveaux hiérarchisés pour estimer la quantité d'intérêt à l'aide du modèle le moins coûteux, puis corrigent successivement cette estimation à l'aide des modèles plus précis. Dans certains cas spécifiques profitant bien à différentes méthodes, il est possible de combiner les dites méthodes afin de profiter des avantages de chacune comme avec les variables de contrôle multi-niveaux [22, 1] ou le quasi-Monte Carlo multi-niveaux [26]. Plus récemment, l'introduction du meilleur estimateur linéaire non-biaisé multi-niveaux (MLBLUE) [63, 62, 14, 18] permet une généralisation de différents estimateurs et notamment du MLMC en combinant les différents niveaux de manière optimale pour minimiser la variance. Parmi ces différentes méthodes, le MLMC est particulièrement adapté pour des applications en géosciences où l'on a accès à une collection de simulateurs basés sur des grilles de différentes fidélités et c'est donc la méthode étudiée dans ce manuscrit.

L'estimateur MLMC, introduit par Heinrich en 1998 [31, 32], a été fortement popularisé par Giles une dizaine d'années plus tard [24, 23]. Il permet d'obtenir une estimation de l'espérance d'une quantité d'intérêt avec la même MSE qu'un estimateur MC mais pour un coût de calcul plus faible. Il appartient à la catégorie des techniques de réduction de variance, c'est-à-dire qu'il permet de réduire la variance d'un estimateur sans détériorer son biais comparé à un simple estimateur MC. D'abord utilisé pour résoudre des équations aux dérivées partielles et des équations différentielles stochastiques [12, 67, 13], des preuves de convergence pour l'estimation de moments statistiques d'ordre arbitraire [3, 4] et de mesures de sensibilité [52] permettent d'élargir ses cas d'application. Plusieurs travaux ont également permis son utilisation pour des quantités appartenant à des espaces de Hilbert [3] et de Banach [38, 37]. Des études sur la combinaison des méthodes multigrilles avec l'estimateur MLMC [39] tirent profit de la collection de modèles utilisés qui convient aux deux méthodes pour des gains supplémentaires au sein de certaines applications. Ces extensions permettent donc d'utiliser le MLMC dans des domaines divers avec, de manière non-exhaustive, des applications en finance [24, 25], au transport de fluides [19], aux modélisations de réservoir [27] ou pour des systèmes biologiques [45].

Le problème étudié dans ce manuscrit est celui de l'estimation des coefficients de normalisation d'une matrice de covariance obtenue à l'aide d'un opérateur de diffusion [73]. Ce problème apparaît notamment en assimilation de données pour l'océan [70], l'atmosphère [46] ou en géostatistique [58]. Le point de départ est l'utilisation d'un opérateur de diffusion pour construire des matrices de covariance avec

la structure souhaitée, et l'étape suivante consiste à fixer les variances de la matrice obtenue, ce qui requiert d'abord de la normaliser. La normalisation d'une matrice de covariance repose sur l'estimation de ces éléments diagonaux, estimation qui se base, notamment, sur de la randomisation avec un estimateur Monte Carlo de l'espérance d'un champ discrétisé aléatoire. L'ensemble des champs discrétisés utilisé est généré à partir d'un modèle, basé sur l'opérateur discrétisé par différence finie de la diffusion, que l'on applique à des champs aléatoires en entrée, une étape coûteuse qui limite la précision atteignable pour les coefficients de normalisation. Il est cependant facile de créer des versions basse fidélité d'un tel modèle en utilisant une discrétisation plus grossière de l'espace lors de la construction de l'opérateur de diffusion. L'estimateur MLMC semble donc être un bon candidat pour obtenir les coefficients de normalisation avec une erreur quadratique moyenne plus faible pour le même coût de calcul que l'estimateur Monte Carlo utilisé. Les spécificités d'une telle application résident, premièrement, dans le fait que les entrées et sorties des simulateurs sont des champs discrétisés. Deuxièmement, dans l'opérateur de diffusion qui est construit à partir d'une approximation discrète de la formulation forte de l'équation de diffusion avec des différences finies, et donc, contrairement à une résolution par éléments finis, il n'y a pas de cadre fonctionnel.

On insiste ici sur la notion de champs discrétisés qui sera utilisée abondamment dans la suite. Un champ discrétisé est couramment représenté comme un vecteur dont chaque entrée est associée à la valeur du champ à un point de l'espace. Une relation physique relie donc les entrées du vecteur qui ne sont pas à considérer comme des valeurs indépendantes. Un champ est décomposable en composantes associées à différentes échelles (ou fréquences) spatiales qui peuvent correspondre à des phénomènes physiques différents. Cette décomposition est importante car l'utilisation de grilles de différentes résolutions dans les modèles conduit à des représentations très différentes de ces échelles. Comprendre les différences de traitement des différentes échelles lors d'une estimation est donc primordial pour savoir lesquelles sont bien représentées dans le résultat final, notamment si la quantité d'intérêt est elle aussi un champ discrétisé et pas seulement une quantité scalaire. Pour garantir l'efficacité de l'estimateur MLMC, un couplage entre les sorties des modèles est nécessaire et ces sorties doivent donc être de dimensions compatibles. Le respect de ces contraintes implique des adaptations dont on souhaite connaître les effets sur les composantes fréquentielles des champs considérés.

L'impact de grilles de différentes résolutions sur les composantes d'échelle est connu depuis plusieurs années pour les méthodes multigrilles [7, 69, 76] et commence à être abordé pour le MLMC. Les travaux de Croci [15] en font partie et portent sur l'échantillonnage de bruit blanc à partir de modèles basés sur les éléments finis. Les éléments finis permettent, cependant, de rester dans un cadre fonctionnel, ce qui n'est pas le cas du problème que l'on considère qui se base sur des différences finies. On mentionnera également les récents papiers de Istratuca [35] et Schwab [64] qui, comme nous, utilisent des champs discrétisés comme entrées de leurs modèles, mais considèrent une quantité d'intérêt scalaire. Ils évoquent de potentiels problèmes causés par les composantes hautes fréquences (ou petites échelles) des champs d'entrées et proposent donc de tronquer leur représentation [64] ou d'ajouter une étape de filtrage [35]. Le passage d'un estimateur MLMC pour une quantité scalaire à un estimateur pour des champs discrétisés n'est pas trivial, d'autant plus sans cadre fonctionnel, et les adaptations nécessaires peuvent apporter de nouveaux problèmes sans une bonne compréhension de leurs effets. Le problème choisi ici se base sur des simulateurs dont les entrées et sorties sont des champs pour lesquels seule une représentation discrète est accessible et l'utilisation de différentes grilles implique

des représentations hétérogènes de ces champs à travers les différents niveaux. Ce contexte impose donc des adaptations de l'estimateur MLMC dont les effets sur les champs considérés doivent être étudiés afin d'assurer son efficacité.

Les contributions de cette thèse sont les suivantes :

- L'adaptation de l'estimateur MLMC de l'espérance au cas spécifique où la quantité d'intérêt est la sortie d'un simulateur appliqué à des vecteurs aléatoires, ces vecteurs aléatoires représentant des champs discrétisés. Autrement dit, on cherche à estimer  $\mathbb{E}[f(\mathbf{X})]$  où  $\mathbf{X} \in L^2(\Omega, \mathbb{R}^n)$  et  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Cette estimation multi-niveaux utilise une hiérarchie de grilles sur laquelle les entrées et sorties des modèles ont des dimensions dépendantes de la discrétisation. L'estimateur requiert donc l'introduction d'opérateurs de transfert de grilles : un opérateur de restriction permettant de transférer un champ discrétisé d'une grille fine vers une grille plus grossière et un opérateur de prolongation pour le transfert d'une grille grossière vers une grille fine (cf. chapitre 3).
- Une analyse spectrale de l'estimateur MLMC basé sur des grilles centrées sur les cellules est proposée afin d'étudier l'effet de ces opérateurs de transfert sur les différentes composantes d'échelle de la quantité d'intérêt. Cette analyse comprend d'abord une partie théorique ayant pour objectif la compréhension des effets des opérateurs de transfert. Cette partie analytique est proposée dans le cadre spécifique d'un estimateur MLMC bi-niveaux impliquant un simulateur linéaire, symétrique et circulant. L'obtention d'une expression de la variance dans ce cas spécifique permet de quantifier la mauvaise estimation de certaines fréquences qui peut advenir pour certains opérateurs de transferts. Une expérience numérique sur une illustration 1D simple est ensuite présentée pour vérifier les résultats théoriques obtenus. La variance de l'estimateur MLMC sur cette illustration est décomposée dans un espace spectral et comparée à celle d'un simple estimateur MC. Cette décomposition permet de constater l'impact négatif que les opérateurs de transferts peuvent avoir sur l'estimation des hautes fréquences (cf. chapitre 3).
- L'introduction de l'estimateur filtered-MLMC (FMLMC) qui ajoute un opérateur de filtrage des hautes fréquences avant l'application de la restriction et après l'application de la prolongation. L'ajout du filtrage peut aussi être vu comme une modification des opérateurs de transfert de grilles dont on augmente l'ordre. L'analyse spectrale de ces opérateurs de transfert modifiés permet de déduire différentes propriétés souhaitées pour éviter les effets observés plus tôt. L'estimateur weighted MLMC (WMLMC) est ici utilisé afin de trouver les opérateurs de transfert optimaux dans certains cas spécifiques (cf. chapitre 4).
- L'application de l'estimateur FMLMC pour l'estimation des coefficients de normalisation d'un opérateur de diffusion en 2D pour des champs de diffusivité hétérogène. Les résultats montrent une réduction de l'erreur quadratique moyenne pour un budget fixé d'un tel estimateur comparé à un MLMC ou un simple MC (cf. chapitre 5).

Le manuscrit est organisé de la manière suivante. Le chapitre 2 rappelle les fondamentaux de l'estimateur MLMC ainsi que ses différentes propriétés. L'adaptation de l'estimateur aux champs discrétisés sur des grilles centrées sur les cellules est ensuite présentée dans le chapitre 3, ainsi que son analyse spectrale dans un cadre spécifique. Cette analyse conduit à l'introduction de l'estimateur filtered-MLMC dans le chapitre 4. Ce chapitre présente également une étude du choix des différents opérateurs de transfert et de filtrage utilisés. Les performances de l'estimateur FMLMC sont illustrées avec son application au problème de l'estimation des coefficients de

normalisation dans le chapitre 5 avant de conclure et de donner des perspectives pour de futurs travaux dans le chapitre 6.

## Chapitre 2

# Monte Carlo multi-niveaux

### 2.1 Notations

Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace probabilisé avec  $\Omega$  l'espace des possibles,  $\mathcal{F}$  la tribu des sous-ensembles de l'espace des possibles, et  $\mathbb{P}$  une mesure de probabilité. Dans ce manuscrit on considère un espace de Hilbert  $H$  muni du produit scalaire  $\langle \cdot, \cdot \rangle_H$  et de la norme associée  $\| \cdot \|_H$ . On se restreint à l'étude de variables aléatoires à valeur dans  $H$  et de moment d'ordre 2 fini appartenant donc à l'espace

$$L^2(\Omega, H) := \{ \xi : \Omega \rightarrow H \mid \int_{\Omega} \|\xi(\omega)\|_H^2 d\mathbb{P}(\omega) < +\infty \}. \quad (2.1)$$

L'espace  $L^2(\Omega, H)$  est un espace de Hilbert quand il est muni du produit scalaire  $\langle \cdot, \cdot \rangle_{L^2(\Omega, H)}$  défini comme

$$\forall \xi, \eta \in L^2(\Omega, H), \quad \langle \xi, \eta \rangle_{L^2(\Omega, H)} := \int_{\Omega} \langle \xi(\omega), \eta(\omega) \rangle_H d\mathbb{P}(\omega). \quad (2.2)$$

Ce produit scalaire induit la norme  $\| \cdot \|_{L^2(\Omega, H)}$ . Par la suite on utilisera la notation  $\mathbb{E}[\| \cdot \|_H^2]^{1/2} = \| \cdot \|_{L^2(\Omega, H)}$ . L'espace de Hilbert  $H$  considéré changera selon la section du manuscrit mais sera précisé à chaque fois.

### 2.2 Monte Carlo pour l'espérance d'une variable aléatoire réelle

On commence par rappeler le principe et les propriétés de l'estimateur Monte Carlo [61, 44]. On considère le cas de variables aléatoires à valeur dans  $\mathbb{R}$  avec un modèle numérique  $f : \mathbb{R} \rightarrow \mathbb{R}$ . L'espace de Hilbert  $H$  considéré est donc  $\mathbb{R}$  muni de la multiplication. On cherche à estimer l'espérance de la variable aléatoire  $f(X) \in L^2(\Omega, \mathbb{R})$  pour  $X \in L^2(\Omega, \mathbb{R})$ , donc telle que  $\mathbb{V}[f(X)] < +\infty$ . L'estimateur Monte Carlo de  $\mu := \mathbb{E}[f(X)]$  est

$$\hat{\mu} = \frac{1}{M} \sum_{i=1}^M f(X^{(i)}) \quad (2.3)$$

où  $\mathcal{X} = \{X^{(i)}\}_{i=1}^M$  est un  $M$ -échantillon de  $X$ , c'est-à-dire une famille de  $M$  variables aléatoires indépendantes et identiquement distribuées selon la même distribution que  $X$ . Cet estimateur est non-biaisé, en effet

$$\mathbb{E}[\hat{\mu}] = \frac{1}{M} \sum_{i=1}^M \mathbb{E}[f(X^{(i)})] = \mathbb{E}[f(X)]. \quad (2.4)$$

La variance de cet estimateur peut être exprimée en utilisant l'indépendance des variables aléatoires du  $M$ -échantillon de  $X$  :

$$\mathbb{V}[\hat{\mu}] = \frac{1}{M^2} \sum_{i=1}^M \mathbb{V}[f(X^{(i)})] = \frac{\mathbb{V}[f(X)]}{M}. \quad (2.5)$$

Pour mesurer la précision d'un estimateur  $\hat{\theta}$  d'une statistique  $\theta$  on utilise la racine de l'erreur quadratique moyenne, root mean square error en anglais (RMSE), définie comme la racine carrée de l'erreur quadratique moyenne (MSE). La MSE est définie comme suit

$$\text{MSE}(\hat{\theta}, \theta) := \mathbb{E}[(\hat{\theta} - \theta)^2]. \quad (2.6)$$

Cette MSE peut être écrite comme la somme de la variance de l'estimateur et de son biais au carré,

$$\text{MSE}(\hat{\theta}, \theta) = \mathbb{V}[\hat{\theta}] + \text{Biais}(\hat{\theta}, \theta)^2 \quad (2.7)$$

avec  $\text{Biais}(\hat{\theta}, \theta) = \mathbb{E}[\hat{\theta}] - \theta$ . Dans le cas de l'estimateur Monte Carlo  $\hat{\mu}$  de  $\mathbb{E}[f(X)]$  cela donne

$$\text{MSE}(\hat{\mu}, \mathbb{E}[f(X)]) = \frac{\mathbb{V}[f(X)]}{M}. \quad (2.8)$$

La RMSE décroît en  $\mathcal{O}(M^{-1/2})$ , c'est-à-dire que pour réduire l'erreur de l'estimateur d'un facteur  $r$  il faut augmenter la taille de l'échantillon par un facteur  $r^2$ . Ce taux de convergence de l'estimateur MC devient rapidement un frein quand le coût de génération des échantillons est élevé.

### 2.3 MLMC pour l'espérance d'une variable aléatoire réelle

On reste, ici, dans le cas de variables aléatoires à valeur dans  $\mathbb{R}$  et on souhaite estimer l'espérance d'une variable aléatoire  $Y \in L^2(\Omega, \mathbb{R})$ . On suppose que cette variable aléatoire n'est pas directement échantillonnable mais qu'elle peut être approximée comme la sortie d'un simulateur appliqué à une variable aléatoire  $X \in L^2(\Omega, \mathbb{R})$ . On considère une hiérarchie multi-niveaux de  $L + 1$  simulateurs numériques  $(f_\ell: \mathbb{R} \rightarrow \mathbb{R})_{\ell=0}^L$  tels que  $f_\ell(X)$  approxime  $Y$  avec une fidélité qui croît avec  $\ell$ . Plus précisément, on suppose que la quantité  $|\mathbb{E}[f_\ell(X) - Y]|$  décroît quand  $\ell$  croît. Le simulateur  $f_0$  est le plus grossier, et donc le moins précis, tandis que le simulateur  $f_L$  est le plus fin, le plus précis. En pratique, le coût moyen nécessaire à l'application d'un simulateur  $f_\ell$  augmente avec  $\ell$ . On présente ici un résumé de l'estimateur tel qu'introduit par Giles [24, 23]. Étant donnée une suite  $(\{X^{(i,\ell)}\}_{i=1}^{M_\ell})_{\ell=0}^L$  de  $L + 1$  échantillons de  $X$  indépendants, l'estimateur Monte Carlo multi-niveaux (MLMC) de  $\mu := \mathbb{E}[Y]$  est

$$\hat{\mu}_L^{\text{MLMC}} = \frac{1}{M_0} \sum_{i=1}^{M_0} f_0(X^{(i,0)}) + \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} (f_\ell(X^{(i,\ell)}) - f_{\ell-1}(X^{(i,\ell)})). \quad (2.9)$$

Le biais d'un estimateur MLMC est le biais du niveau de fidélité le plus fin grâce à la somme télescopique

$$\mathbb{E}[\hat{\mu}_L^{\text{MLMC}}] = \mathbb{E}[f_0(X)] + \sum_{\ell=1}^L (\mathbb{E}[f_\ell(X)] - \mathbb{E}[f_{\ell-1}(X)]) = \mathbb{E}[f_L(X)]. \quad (2.10)$$

Dans le cas de l'estimateur  $\hat{\mu}_L^{\text{MLMC}}$  de  $\mathbb{E}[Y]$ , cela donne donc

$$\text{Biais}(\hat{\mu}_L^{\text{MLMC}}, \mu) = \mathbb{E}[\hat{\mu}_L^{\text{MLMC}}] - \mathbb{E}[Y] = \mathbb{E}[f_L(X)] - \mathbb{E}[Y] = \text{Biais}(f_L(X), Y). \quad (2.11)$$

La variance de l'estimateur  $\hat{\mu}_L^{\text{MLMC}}$  est

$$\mathbb{V}[\hat{\mu}_L^{\text{MLMC}}] = \frac{\mathbb{V}[f_0(X)]}{M_0} + \sum_{\ell=1}^L \frac{\mathbb{V}[f_\ell(X) - f_{\ell-1}(X)]}{M_\ell}. \quad (2.12)$$

En utilisant la décomposition de la MSE de l'équation (2.7), on a

$$\text{MSE}(\hat{\mu}_L^{\text{MLMC}}, \mu) = \frac{\mathbb{V}[f_0(X)]}{M_0} + \sum_{\ell=1}^L \frac{\mathbb{V}[f_\ell(X) - f_{\ell-1}(X)]}{M_\ell} + (\mathbb{E}[f_L(X)] - \mathbb{E}[Y])^2. \quad (2.13)$$

À l'aide de la supposition comme quoi  $|\mathbb{E}[f_\ell(X) - Y]|$  décroît avec  $\ell$ , il est possible de réduire le biais en ajoutant un niveau de fidélité plus fin. Le terme de variance, lui, correspond à l'erreur d'échantillonnage et peut être réduit en augmentant les tailles  $M_\ell$  des ensembles. Il peut également être réduit en ayant des faibles variances pour les termes de corrections, termes qui peuvent être décomposés comme suit

$$\mathbb{V}[f_\ell(X) - f_{\ell-1}(X)] = \mathbb{V}[f_\ell(X)] + \mathbb{V}[f_{\ell-1}(X)] - 2\mathbf{C}[f_\ell(X), f_{\ell-1}(X)]. \quad (2.14)$$

En ayant des approximations de  $Y$  sur différents niveaux de fidélités corrélées positivement entre elles on peut attendre une réduction de la variance de la correction entre deux approximations de  $Y$ . Maximiser la corrélation entre les différents niveaux de fidélité que l'on souhaite utiliser est primordial pour atteindre une réduction de variance intéressante. Par la suite on notera la valeur absolue de l'espérance et la variance d'un terme de correction avec

$$\mathcal{E}_\ell := |\mathbb{E}[f_\ell(X) - f_{\ell-1}(X)]| \quad \text{et} \quad \mathcal{V}_\ell := \mathbb{V}[f_\ell(X) - f_{\ell-1}(X)] \quad (2.15)$$

pour  $\ell = 1, \dots, L$  et  $\mathcal{E}_0 := \mathbb{E}[f_0(X)]$ ,  $\mathcal{V}_0 := \mathbb{V}[f_0(X)]$ .

Le théorème suivant, présenté dans [23], fournit la vitesse de croissance asymptotique de l'estimateur MLMC en fonction de sa précision, exprimée en termes de RMSE.

**Théorème 1.** *Soit  $Y$  une variable aléatoire et  $(f_\ell(X))_{\ell \geq 0}$  une suite de variables aléatoires réelles et de variance finie telle que  $\lim_{\ell \rightarrow +\infty} \mathbb{E}[f_\ell(X)] = \mathbb{E}[Y]$ . Soit  $(\mathcal{C}_\ell)_{\ell \geq 0}$  une suite de réels représentant le coût de calcul moyen nécessaire à la génération d'une réalisation de  $f_\ell(X)$ . Soit  $(n_\ell)_{\ell \geq 0}$  une suite d'entiers telle que  $n_\ell \approx s^\ell$  pour  $s > 1$  fixé. Supposons qu'il existe des constantes  $\alpha, \beta, \gamma > 0$  telles que  $\min(\beta, \gamma) < 2\alpha$ , et, pour  $\ell \geq 0$ ,*

$$(i) \quad |\mathbb{E}[f_\ell(X) - Y]| \lesssim n_\ell^{-\alpha},$$

$$(ii) \quad \mathcal{V}_\ell \lesssim n_\ell^{-\beta},$$

$$(iii) \quad \mathcal{C}_\ell \lesssim n_\ell^\gamma.$$

Alors, pour tout  $\varepsilon < e^{-1}$ , il existe un entier naturel  $L$  et une suite d'entiers naturels  $\{M_\ell\}_{\ell=0}^L$  tels que

$$\text{MSE}(\hat{\mu}_L^{\text{MLMC}}, \mathbb{E}[Y]) < \varepsilon^2 \quad (2.16)$$

avec un coût de calcul moyen

$$\text{Coût}(\hat{\mu}_L^{\text{MLMC}}) \lesssim \begin{cases} \varepsilon^{-2}, & \beta > \gamma, \\ \varepsilon^{-2}(\log \varepsilon)^2, & \beta = \gamma, \\ \varepsilon^{-2-\frac{\gamma-\beta}{\alpha}}, & \beta < \gamma. \end{cases} \quad (2.17)$$

Dans ce théorème, le symbole  $\lesssim$  indique que, pour  $a, b > 0$ ,  $a \lesssim b$  équivaut à dire que  $a/b$  est majoré par une constante, tandis que  $a \approx b$  équivaut à  $a \lesssim b$  et  $b \lesssim a$ . L'hypothèse (i) indique que le biais, en valeur absolue, des approximations de  $Y$  doit décroître exponentiellement à un taux  $\alpha$  quand  $\ell$  croît. L'hypothèse (ii) concerne les variances des termes de corrections qui doivent décroître exponentiellement avec  $\ell$  à un taux  $\beta$ . La dernière hypothèse porte sur le coût de calcul moyen  $\mathcal{C}_\ell$  pour générer une réalisation de  $f_\ell(X)$ , qui doit croître exponentiellement avec un taux  $\gamma$  à travers les niveaux. La majoration du coût de calcul de l'estimateur MLMC dépend du rapport entre  $\beta$  et  $\gamma$ , et la vitesse de croissance du coût la plus favorable correspond au cas  $\beta > \gamma$ . On aimerait donc avoir les variances des termes de correction qui décroissent plus vite que le coût de génération des réalisations ne croît.

En comparaison, un estimateur Monte Carlo  $\hat{\mu}_L$  de  $\mathbb{E}[Y]$  se basant sur le modèle fin  $f_L$  possède la MSE suivante :

$$\text{MSE}(\hat{\mu}_L, \mathbb{E}[Y]) = \frac{\mathbb{V}[f_L(X)]}{M} + \mathbb{E}[f_L(X) - Y]^2. \quad (2.18)$$

Sous les mêmes hypothèses que celles du théorème 1, on a

$$|\mathbb{E}[f_L(X) - Y]| \lesssim n_L^{-\alpha} \quad \text{et} \quad \mathcal{C}_L \lesssim n_L^\gamma. \quad (2.19)$$

En visant

$$\text{MSE}(\hat{\mu}_L, \mathbb{E}[Y]) < \varepsilon^2 \quad (2.20)$$

on cherche à avoir les deux termes de la MSE inférieurs à  $(1/2)\varepsilon^2$ . En notant  $a \sim b$  une relation de proportionnalité entre  $a$  et  $b$ , on obtient  $n_L \sim \varepsilon^{-\frac{1}{\alpha}}$  et  $M \sim \varepsilon^{-2}$  en remarquant que  $\mathbb{V}[f_L(X)]$  est constante et indépendante de  $M$ . On a finalement  $\mathcal{C}_L \sim \varepsilon^{-\frac{\gamma}{\alpha}}$  et

$$\text{Coût}(\hat{\mu}_L) = \mathcal{C}_L M \lesssim \varepsilon^{-2-\frac{\gamma}{\alpha}}. \quad (2.21)$$

Selon le théorème 1, le coût de calcul moyen d'un estimateur MLMC, dans le cas optimal  $\beta > \gamma$ , est inférieur à celui d'un MC d'un facteur  $\varepsilon^{-\frac{\gamma}{\alpha}}$ .

Les tailles d'ensembles  $\{M_\ell\}_{\ell=0}^L$  s'obtiennent en minimisant la variance de l'estimateur MLMC (2.12) pour un coût fixé. C'est ce qui est fait, par exemple, dans [23, 12] et qui donne, en cherchant à obtenir une variance égale à  $\varepsilon^2$ ,

$$M_\ell = \frac{\varepsilon^{-2}}{S_L} \sqrt{\frac{\mathcal{V}_\ell}{\mathcal{C}_\ell + \mathcal{C}_{\ell-1}}}, \quad S_L := \sum_{\ell=1}^L \sqrt{\mathcal{V}_\ell(\mathcal{C}_\ell + \mathcal{C}_{\ell-1})}, \quad (2.22)$$

avec la convention  $\mathcal{C}_{-1} = 0$ . Dans cette expression, les variance  $\mathcal{V}_\ell$  de la correction entre  $f_\ell(X)$  et  $f_{\ell-1}(X)$  sont inconnues et doivent donc être estimées.

Dans la formulation classique de l'estimateur MLMC [23] on considère que le niveau de fidélité le plus grossier est le niveau initial utilisé dans l'estimation. On ajoute ensuite des niveaux plus fins, au besoin, pour réduire la MSE. En pratique, il peut être plus utile de considérer l'inverse. Le niveau initial est celui de plus haute fidélité et on ajoute des niveaux plus grossiers, au besoin, pour réduire la

MSE. Cette vision est plus adaptée aux cas où il existe un modèle fin pour lequel un raffinement supplémentaire n'est pas accessible avec les moyens actuels, mais où des modèles grossiers existent déjà ou sont faciles à construire. De plus, s'il est courant de viser une certaine précision que l'on cherchera à atteindre à moindre coût, il est parfois plus pertinent de chercher à minimiser l'erreur de l'estimateur pour un budget fixé. L'utilisation d'un estimateur MLMC sous la contrainte d'un budget maximum change notamment la manière de calculer les tailles des échantillons  $\{M_\ell\}_{\ell=0}^L$ . Pour un budget  $\eta$  maximum imposé, [52] donne la formule suivante pour  $\ell = 0, \dots, L$

$$M_\ell = \frac{\eta}{S_L} \sqrt{\frac{\mathcal{V}_\ell}{C_\ell + C_{\ell-1}}}. \quad (2.23)$$

Cette formule requiert, comme précédemment, l'estimation des variances  $\mathcal{V}_\ell$  de la correction entre  $f_\ell(X)$  et  $f_{\ell-1}(X)$ .

## 2.4 Illustration MC et MLMC

### 2.4.1 Présentation de l'illustration

Afin d'illustrer les différences entre un estimateur MC et un estimateur MLMC dans un cas scalaire, on reprend, en partie, le problème étudié dans [52],

$$\begin{cases} \frac{\partial u}{\partial t}(t) = \lambda u(t) \\ u(0) = u_0 \end{cases} \quad (2.24)$$

où  $t \in [0, 1]$ ,  $u: \mathbb{R} \rightarrow \mathbb{R}$ ,  $u_0, \lambda \in \mathbb{R}$ . La solution exacte de ce problème est  $u(t) = u_0 e^{\lambda t}$ ,  $\forall t \in [0, 1]$ . On définit

$$F: \mathbb{R}^2 \times [0, 1] \rightarrow \mathbb{R} \quad (2.25)$$

$$(u_0, \lambda, t) \mapsto u_0 e^{\lambda t}. \quad (2.26)$$

On considère maintenant la solution pour  $t$  fixé à 1 et on pose

$$U_0 \sim \mathcal{N}(\mu_0, \sigma_0^2) \quad \text{et} \quad \Lambda \sim \mathcal{N}(\mu, \sigma^2), \quad (2.27)$$

deux variables aléatoires indépendantes qui suivent une loi normale de moyenne  $\mu_0$  et d'écart-type  $\sigma_0$  (respectivement  $\mu$  et  $\sigma$ ). En notant

$$f: \mathbb{R}^2 \rightarrow \mathbb{R} \quad (2.28)$$

$$(u_0, \lambda) \mapsto F(u_0, \lambda, t = 1), \quad (2.29)$$

on peut définir la variable aléatoire  $Y = f(U_0, \Lambda)$  dont on cherche à estimer l'espérance. Pour ce faire, on construit des simulateurs numériques de différentes fidélités

$$f_\ell: \mathbb{R}^2 \rightarrow \mathbb{R} \quad (2.30)$$

$$(u_0, \lambda) \mapsto u_\ell^{n_\ell}(u_0, \lambda), \quad (2.31)$$

où

$$u_\ell^0: \mathbb{R}^2 \rightarrow \mathbb{R} \quad (2.32)$$

$$(u_0, \lambda) \mapsto u_0, \quad (2.33)$$

et

$$u_\ell^k: \mathbb{R}^2 \rightarrow \mathbb{R} \quad (2.34)$$

$$(u_0, \lambda) \mapsto \frac{u_\ell^{k-1}(u_0, \lambda)}{1 - \lambda/n_\ell} \quad (2.35)$$

pour  $k = 1, \dots, n_\ell$ . Les simulateurs  $f_\ell$  s'apparentent donc à une résolution de l'équation (2.24) par un schéma d'Euler implicite avec un pas  $1/n_\ell$ . On considère que le coût d'une évaluation de  $f_\ell$  est proportionnelle au paramètre de discrétisation  $n_\ell$ , i.e.,  $\mathcal{C}_\ell \sim n_\ell$  pour  $\ell = 0, \dots, L$ .

## 2.4.2 Illustration MC

On définit  $Y_\ell = f_\ell(U_0, \Lambda)$ . On peut utiliser un estimateur MC se basant sur le simulateur  $f_\ell$  pour avoir une estimation de l'espérance de  $Y$ . Pour les expérimentations qui suivent, on considère 6 niveaux de fidélités différents avec  $n_0 = 16$  et  $n_\ell = n_0 2^\ell$  pour  $\ell = 1, \dots, 5$ . Étant donné un  $M_\ell$ -échantillon aléatoire  $\{(U_0^{(i)}, \Lambda^{(i)})\}_{i=1}^{M_\ell}$ , on peut construire un estimateur Monte Carlo de  $\mathbb{E}[Y]$  se basant sur le modèle  $f_\ell$

$$\hat{\mu}_\ell = \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} f_\ell(U_0^{(i)}, \Lambda^{(i)}), \quad (2.36)$$

pour  $\ell = 0, \dots, 5$ . La figure 2.1 montre l'évolution de la MSE des estimateurs  $\hat{\mu}_\ell$  pour  $\ell = 0, \dots, 5$  en fonction de  $M_\ell$ , la taille de l'échantillon.

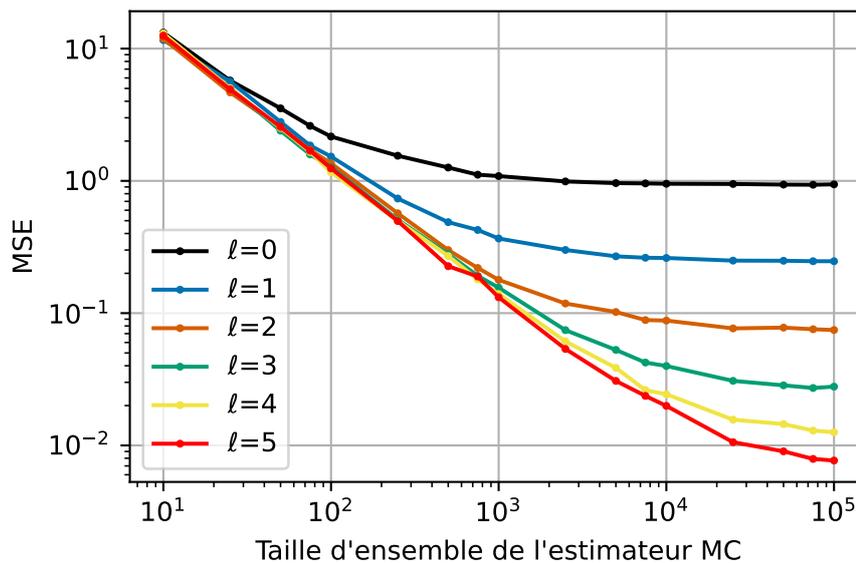


FIGURE 2.1 – Évolution de la MSE de différents estimateurs MC  $\hat{\mu}_\ell$  de  $Y$  se basant sur les modèles  $f_\ell$  en fonction de la taille de l'échantillon  $M_\ell$ , pour  $\ell = 0, \dots, 5$ . La MSE est estimée à partir de 1000 estimateurs.

On observe que la MSE des différents estimateurs semble converger vers une valeur qui dépend du niveau de fidélité de l'estimateur. Cette valeur étant plus faible

pour les niveau de fidélité plus élevés. La MSE peut s'exprimer comme la somme de l'erreur d'échantillonnage et du biais au carré (cf. équation (2.7)) où l'erreur d'échantillonnage décroît en  $1/M_\ell$  et le biais dépend du niveau de fidélité considéré. Ainsi, augmenter la taille de l'échantillon permet de réduire l'erreur d'échantillonnage mais pas le biais qui va finir par dominer la MSE. C'est ce phénomène qui est mis en évidence ici, la MSE des différents estimateurs commence par décroître quand on augmente la taille de l'échantillon car l'erreur d'échantillonnage est de plus en plus faible. Cette décroissance suit bien, pour tous les estimateurs, une pente en  $1/M_\ell$ . La MSE atteint, cependant, un plateau lorsque le biais domine la MSE. Le biais étant plus important sur les niveaux grossiers, la MSE des estimateurs correspondant converge vers une valeur plus élevée. Pour réduire la MSE au delà de ce plateau, il faut donc considérer un niveau plus fin.

Les limites de l'estimateur MC sont, ici, bien visibles. Un niveau de fidélité suffisamment élevé est requis afin de réduire le biais et un échantillon de grande taille permet alors de réduire l'erreur d'échantillonnage. La génération d'un échantillon de grande taille pouvant être coûteuse pour des simulateurs de haute fidélité, un compromis doit être trouvé.

### 2.4.3 Illustration MLMC

On peut, à présent, définir un estimateur MLMC de  $\mathbb{E}[Y_L]$

$$\hat{\mu}_L^{\text{MLMC}} = \frac{1}{M_0} \sum_{i=1}^{M_0} f_0((U_0, \Lambda)^{(i,0)}) + \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \left( f_\ell((U_0, \Lambda)^{(i,\ell)}) - f_{\ell-1}((U_0, \Lambda)^{(i,\ell)}) \right), \quad (2.37)$$

à partir de  $L + 1$  différents échantillons indépendants  $(\{(U_0, \Lambda)^{(i,\ell)}\}_{i=1}^{M_\ell})_{\ell=0}^L$  de  $(U_0, \Lambda)$ .

Le théorème 1 nous assure que, pour une précision fixée (en terme de MSE), le coût de calcul de l'estimateur MLMC est inférieur à celui de l'estimateur MC si les trois hypothèses du théorème sont respectées. On rappelle que l'on considère une évaluation d'un modèle  $f_\ell$ , défini dans l'équation (2.30), comme ayant un coût  $\mathcal{C}_\ell \sim n_\ell$ . La troisième hypothèse du théorème est donc vérifiée et on a  $\gamma = 1$ . On vérifie numériquement la validité des deux autres hypothèses en estimant les quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  de l'équation (2.15) pour  $\ell = 0, \dots, L$  et  $L = 5$ . Les résultats sont présentés dans la figure 2.2.

L'espérance et la variance des termes de correction décroît bien exponentiellement avec la discrétisation  $n_\ell$ . On a  $\min(\beta, \gamma) < 2\alpha$ ,  $\mathcal{E}_\ell \sim n_\ell^{-\alpha}$  et  $\mathcal{V}_\ell \sim n_\ell^{-\beta}$  où  $\alpha \approx 1$  et  $\beta \approx 2$ . Les hypothèses du théorème 1 sont donc bien vérifiées pour cette illustration et l'on se retrouve dans le cas favorable au MLMC où  $\beta > \gamma$ . On s'attend donc à des estimateurs MLMC donc le coût de calcul est inférieur à celui d'un estimateur MC pour la même précision. Ou, dans notre cas, des estimateurs MLMC atteignant des valeurs de MSE plus faibles qu'un estimateur MC pour le même budget  $\eta$ . La figure 2.3 compare l'allocation optimale, obtenue à partir de l'équation (2.23), de différents estimateurs MLMC comprenant un nombre différent de niveaux ainsi que leur MSE. Pour tous ces estimateurs, le niveau le plus fin correspond à  $n_L = 512$ .

La figure 2.3a présente la répartition des échantillons sur les différents niveaux pour un budget fixé équivalent au coût nécessaire pour générer 1000 membres sur le niveau le plus fin. On observe que la taille de l'ensemble du niveau le plus fin est très réduite, ce qui laisse une grande partie du budget total  $\eta$  disponible pour générer des ensembles de grande taille sur les niveaux plus grossiers. Pour le MLMC à 2 niveaux ( $L = 1$ ), un unique membre est gardé sur le niveau fin pour en avoir presque 2000

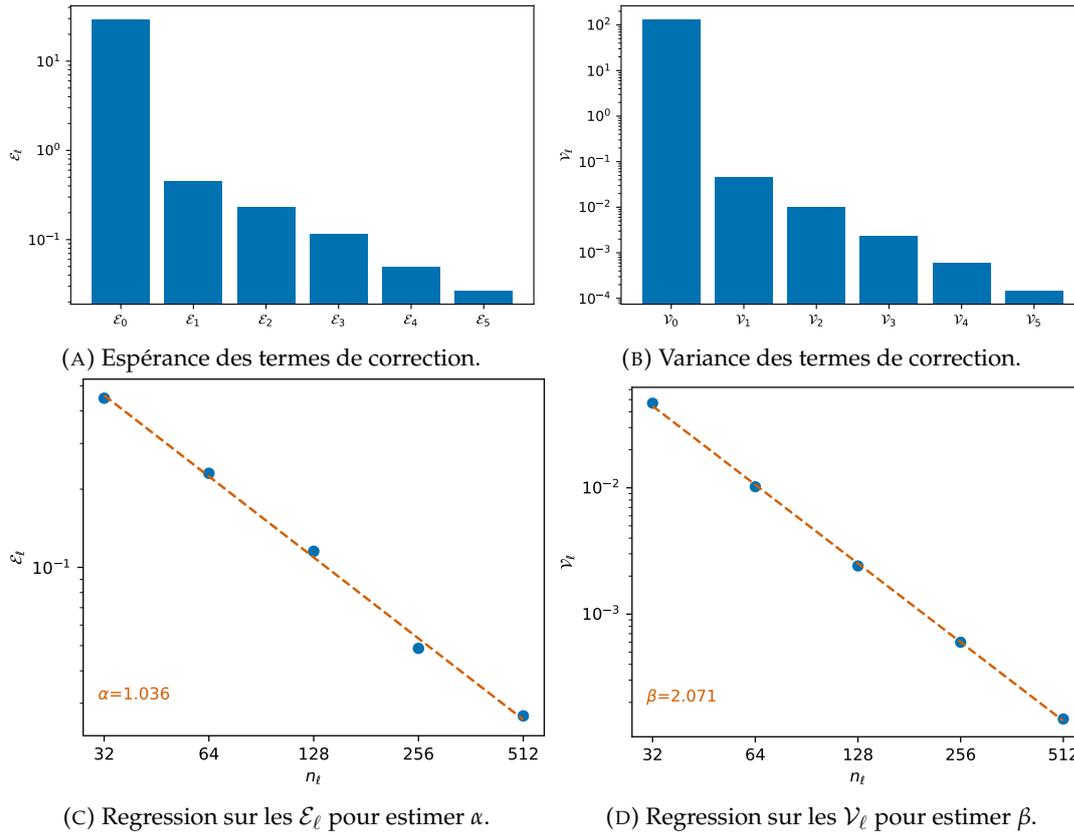


FIGURE 2.2 – Estimation des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  pour  $\ell = 1, \dots, L$  (haut) et regression sur ces quantités afin d'estimer  $\alpha$  et  $\beta$  (bas). Les estimations sont faites à partir d'un échantillon de taille 1000.

sur le niveau grossier. Le MLMC à 6 niveaux atteint une taille d'ensemble proche de 30 000 pour le niveau le plus grossier. La figure 2.3b montre que de telles allocations permettent effectivement de réduire la MSE d'un estimateur MLMC comparé à un estimateur MC pour le même budget, avec près de 90% de réduction pour le MLMC à 6 niveaux.

## 2.5 Conclusion

L'application du MLMC à l'estimation de l'espérance d'une variable aléatoire est un cas simple qui a largement été étudié [24, 23, 12]. Sous les hypothèses de décroissance exponentielle du biais et de la variance des termes de correction, présentées dans le théorème 1, le MLMC permet d'atteindre une MSE fixée pour un coût inférieur à celui d'un MC. L'illustration de la section 2.4.3 a permis de montrer un exemple d'application du MLMC dans un contexte où le budget total est fixé, ce qui conduit à une réduction de la variance du MLMC comparé à un MC. Cependant, le problème de normalisation d'un opérateur de diffusion, étudié dans ce manuscrit, est basé sur l'estimation de paramètres statistiques de champs discrétisés en espace qui sont les sorties d'un modèle appliqué à des champs aléatoires. Dans un tel contexte, l'utilisation de l'estimateur MLMC requiert plusieurs adaptations. L'objectif du chapitre suivant est le traitement de ce cas d'application et la présentation des adaptations nécessaires.

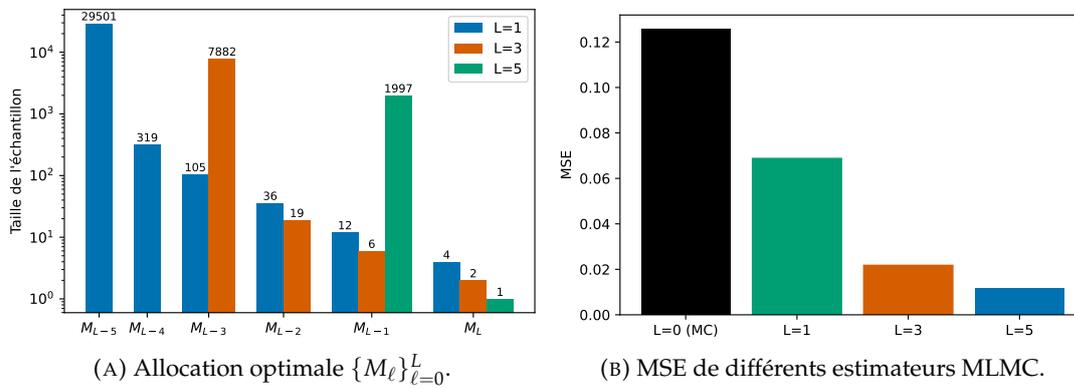


FIGURE 2.3 – Allocation optimale des échantillons sur les différents niveaux et la MSE résultante des estimateurs MC ( $L = 0$ ) et de différents estimateurs MLMC ( $L \in \{1, 3, 5\}$ ). Le niveau le plus fin  $L$  correspond à une résolution  $n_L = 512$  et le budget total est fixé à  $\eta = 1000C_L$ . La MSE est estimée à partir de 10 000 estimateurs.

## Chapitre 3

# MLMC pour des champs discrétisés spatialement

Ce chapitre est consacré à l'adaptation et l'étude de l'estimateur MLMC de l'espérance pour des champs aléatoires discrétisés en espace sur des grilles centrées sur les cellules. On commence par considérer le cas général de champs discrétisés appartenant à des espaces de Hilbert et l'on introduit la notion d'opérateurs de transfert afin de pouvoir construire un estimateur MLMC. On se ramène ensuite à un cas plus spécifique en précisant la discrétisation et les opérateurs de transfert utilisés, motivés par l'application du chapitre 5. Afin d'étudier l'impact des opérateurs de transfert sur l'estimateur MLMC, une analyse spectrale est présentée, et les différents résultats qui en découlent sont illustrés numériquement à l'aide d'un problème jouet unidimensionnel.

### 3.1 MLMC pour des champs discrétisés

On commence par reprendre les notations de la section 2.1 et on considère des champs discrétisés aléatoires de moment d'ordre 2 fini et à valeur dans un espace de Hilbert. Ce contexte est celui considéré dans [3] et on se repose donc sur leurs travaux pour notre étude. Le simulateur haute fidélité est

$$f_L: H_L \rightarrow H_L. \quad (3.1)$$

On note  $X_L \in L^2(\Omega, H_L)$  les champs aléatoires discrétisés d'entrée,  $Y_L = f_L(X_L)$  les champs aléatoires discrétisés de sortie du simulateur  $f_L$  et on suppose également que  $Y_L \in L^2(\Omega, H_L)$ . On cherche à estimer le champ discrétisé  $\mathbb{E}[Y_L] \in H_L$ .

On souhaite utiliser un estimateur MLMC qui, en se basant sur des modèles de basse fidélité approximant  $f_L$ , peut atteindre une variance plus faible qu'un estimateur MC pour le même budget. On considère donc une hiérarchie de modèles  $(f_\ell)_{\ell=0}^{L-1}$  qui sont des approximations du modèle fin  $f_L$ . Dans ce manuscrit, on s'intéresse au cas où les modèles  $f_\ell$  sont construits à partir de différentes discrétisations spatiales et possèdent donc des espaces d'entrée et de sortie dont les dimensions dépendent du niveau  $\ell$ . On pose, pour  $\ell = 0, \dots, L-1$ ,

$$f_\ell: H_\ell \rightarrow H_\ell. \quad (3.2)$$

Les inconsistances dans les dimensions d'entrée et de sortie des différents modèles requiert des adaptations de l'estimateur MLMC. Un tel contexte diffère de celui étudié dans [3] où les modèles considérés ont des espaces d'entrée et de sortie de dimension constante à travers les niveaux.

### 3.1.1 Opérateurs de transfert

On introduit ici les opérateurs de transfert qui servent à transférer un vecteur d'un espace de Hilbert à un autre. Les différents espaces considérés sont les espaces  $H_\ell$  munis du produit scalaire  $\langle \cdot, \cdot \rangle_{H_\ell}$  et de la norme associée  $\| \cdot \|_{H_\ell}$ , définis pour  $\ell = 0, \dots, L$ . On utilisera le terme de restriction pour désigner les opérateurs qui permettent de transférer un champ discrétisé d'un espace de Hilbert de haute fidélité  $H_\ell$  vers l'espace de Hilbert de plus basse fidélité  $H_{\ell-1}$ . De la même manière, on désigne par prolongation les opérateurs permettant de transférer un champ discrétisé de l'espace  $H_{\ell-1}$  vers l'espace  $H_\ell$ . On définit donc, pour  $\ell = 1, \dots, L$ ,

$$R_{\ell-1}^{\ell-1}: H_\ell \rightarrow H_{\ell-1} \quad \text{et} \quad P_{\ell-1}^\ell: H_{\ell-1} \rightarrow H_\ell, \quad (3.3)$$

Avec de tels opérateurs il devient possible de transférer toutes les sorties des modèles dans un même espace afin d'utiliser un estimateur MLMC. Étant donné que l'on souhaite estimer l'espérance de la sortie du modèle haute fidélité, on décide de transférer toutes les sorties des différents niveaux dans  $H_L$ , l'espace de plus haute fidélité. On définit donc

$$P_\ell^L = P_{L-1}^L \circ P_{L-2}^{L-1} \circ \dots \circ P_{\ell+1}^{\ell+2} \circ P_\ell^{\ell+1} \quad (3.4)$$

l'opérateur de prolongation entre l'espace  $H_\ell$  et  $H_L$  pour  $\ell = 0, \dots, L-1$ . Le couplage des entrées des modèles dans les termes de correction de l'estimateur est également primordial et devient possible grâce à ces opérateurs. En effet, pour le terme de correction entre le niveau  $\ell$  et  $\ell-1$ , il est possible de générer un champ discrétisé d'entrée dans l'espace  $H_\ell$  puis de le restreindre dans l'espace  $H_{\ell-1}$  afin de l'utiliser en entrée du modèle  $f_{\ell-1}$ . On définit

$$R_L^\ell = R_{\ell+1}^\ell \circ R_{\ell+2}^{\ell+1} \circ \dots \circ R_{L-1}^{L-2} \circ R_L^{L-1} \quad (3.5)$$

l'opérateur de restriction entre  $H_L$  et  $H_\ell$  pour  $\ell = 0, \dots, L-1$ . On pose  $P_L^L = R_L^L = id$ , la fonction identité. On peut alors construire la suite de modèles de différentes résolutions  $(\tilde{f}_\ell)_{\ell=0}^L$  qui approximent  $f_L$

$$\begin{aligned} \tilde{f}_\ell: H_L &\rightarrow H_L \\ x &\mapsto (P_\ell^L \circ f_\ell \circ R_L^\ell)(x). \end{aligned} \quad (3.6)$$

On suppose que, pour tout  $X_\ell \in L^2(\Omega, H_\ell)$ , on a  $P_{\ell+1}^{\ell+1}(X_\ell) \in L^2(\Omega, H_{\ell+1})$  et  $R_{\ell-1}^{\ell-1}(X_\ell) \in L^2(\Omega, H_{\ell-1})$ , ce qui implique que  $\tilde{Y}_\ell = \tilde{f}_\ell(X_L) \in L^2(\Omega, H_\ell)$  pour  $X_L \in L^2(\Omega, H_L)$ , et cela pour  $\ell = 0, \dots, L$ .

### 3.1.2 Formulation de l'estimateur MLMC pour des champs discrétisés

La suite  $(\tilde{f}_\ell)_{\ell=0}^L$  forme la hiérarchie de modèles qui est utilisée pour l'estimateur MLMC, tous les modèles  $\tilde{f}_\ell$  ont  $H_L$  en espace d'entrée et de sortie. Si le coût d'application des opérateurs de transfert est négligeable devant le coût d'application du modèle  $f_\ell$ , ce que l'on suppose par la suite, on conserve la hiérarchie de coût de calcul de la suite  $(f_\ell)_{\ell=0}^L$ . On peut alors construire un estimateur MLMC de  $\mathbb{E}[Y_L]$

$$\hat{\mu}_L^{\text{MLMC}} = \frac{1}{M_0} \sum_{i=1}^{M_0} \tilde{f}_0(X_L^{(i,0)}) + \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \left( \tilde{f}_\ell(X_L^{(i,\ell)}) - \tilde{f}_{\ell-1}(X_L^{(i,\ell)}) \right), \quad (3.7)$$

à partir de  $L + 1$  différents échantillons indépendants  $(\{X_L^{(i,\ell)}\}_{i=1}^{M_\ell})_{\ell=0}^L$  de  $X_L$ .

L'erreur quadratique moyenne de cet estimateur par rapport au paramètre  $\mu$  s'écrit

$$\text{MSE}(\hat{\mu}_L^{\text{MLMC}}, \mu) = \|\hat{\mu}_L^{\text{MLMC}} - \mu\|_{L^2(\Omega, H_L)}^2 = \mathbb{E}[\|\hat{\mu}_L^{\text{MLMC}} - \mu\|_{H_L}^2]. \quad (3.8)$$

Comme démontré dans [3, Theorem 3.1], cette MSE peut être décomposée en un terme correspondant au biais de discrétisation au carré et un terme égal à l'erreur d'échantillonnage de l'estimateur :

$$\text{MSE}(\hat{\mu}_L^{\text{MLMC}}, \mu) = \mathcal{V}(\hat{\mu}_L^{\text{MLMC}}) + \|\mathbb{E}[\hat{\mu}_L^{\text{MLMC}}] - \mu\|_{H_L}^2, \quad (3.9)$$

où

$$\forall Z \in L^2(\Omega, H_L), \quad \mathcal{V}(Z) := \|Z - \mathbb{E}[Z]\|_{L^2(\Omega, H_L)}^2 = \mathbb{E}[\|Z\|_{H_L}^2] - \|\mathbb{E}[Z]\|_{H_L}^2 \quad (3.10)$$

est analogue à une variance pour un champ discrétisé aléatoire et

$$\mathcal{V}(\hat{\mu}_L^{\text{MLMC}}) = \frac{1}{M_0} \mathcal{V}(\tilde{Y}_0) + \sum_{\ell=1}^L \frac{1}{M_\ell} \mathcal{V}(\tilde{Y}_\ell - \tilde{Y}_{\ell-1}) \quad (3.11)$$

est la variance de l'estimateur MLMC. Dans notre cas, le paramètre étudié est  $\mu_L = \mathbb{E}[Y_L]$ , l'espérance de la sortie du modèle haute fidélité. Étant donné que  $\mathbb{E}[\hat{\mu}_L^{\text{MLMC}}] = \mu_L$  par construction, le biais de notre estimateur par rapport à  $\mu_L$  est nul et on a

$$\text{MSE}(\hat{\mu}_L^{\text{MLMC}}, \mu) = \mathcal{V}(\hat{\mu}_L^{\text{MLMC}}). \quad (3.12)$$

Dans la suite du manuscrit, on étudie donc seulement la variance de l'estimateur, définie avec les équations (3.10) et (3.11), qui correspond à son erreur quadratique moyenne par rapport au paramètre  $\mu_L$ .

## 3.2 Formulation forte et discrétisation

Une différence majeure entre les travaux de ce manuscrit et ceux présentés dans [64, 35, 15], où les modèles ont également des espaces d'entrée et de sortie de dimensions inconsistantes à travers les niveaux, repose sur les discrétisations considérées et la nature des espaces de Hilbert  $H_\ell$ . Dans les travaux cités, les simulateurs considérés sont construits à partir d'une discrétisation de la formulation faible d'une équation aux dérivées partielles (EDP), spécifiquement en utilisant la méthode des éléments finis. Les espaces de Hilbert  $H_\ell$  sont donc des espaces fonctionnels. Cette thèse traite le cas de modèles  $(f_\ell)_{\ell=0}^L$  construits comme une approximation discrète de la formulation forte d'une EDP. Ce choix d'étude provient de l'application présentée dans la section 5.1 qui se repose sur des différences finies pour construire les différents simulateurs. Les champs discrétisés étudiés ne sont donc représentables que par un vecteur dont les composantes correspondent à la valeur du champ à un point de l'espace. Les espaces  $H_\ell$  des équations (3.1) et (3.2) sont les espaces  $\mathbb{R}^{n_\ell}$  où  $n_\ell$  est le nombre de points de discrétisation du domaine spatial considéré. Étant donné que les vecteurs d'entrée et de sortie des différents modèles représentent des champs discrétisés, leurs composantes ne sont pas indépendantes les unes des autres, il existe une relation physique les reliant. L'étude de simulateurs basés sur une formulation faible ou forte d'une EDP a un impact fort sur le choix des discrétisations de l'espace et des opérateurs de transfert considérés, ainsi que sur la manière d'analyser les effets

de ces opérateurs sur l'estimateur MLMC. On présente donc ici la discrétisation de l'espace et les opérateurs de transfert utilisés.

### 3.2.1 Discrétisation de l'espace par grilles

Chaque simulateur  $f_\ell$ , pour  $\ell = 0, \dots, L$ , est basé sur une discrétisation de l'espace par une grille de taille  $n_\ell$  avec  $n_0 < n_1 < \dots < n_L$ . On définit, pour  $\ell = 0, \dots, L$ , la matrice de Gram symétrique, définie positive  $\mathbf{W}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$  qui contient les informations relatives à la représentation d'une fonction continue sur la grille de taille  $n_\ell$  considérée. Il existe, pour  $\ell = 0, \dots, L$ , une matrice  $\mathbf{V}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$  telle que  $\mathbf{W}_\ell$  admette la décomposition  $\mathbf{W}_\ell = \mathbf{V}_\ell \mathbf{V}_\ell^T$ . On peut alors, pour  $\ell = 0, \dots, L$ , définir le produit scalaire

$$\forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^{n_\ell}, \quad \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbf{W}_\ell} = \mathbf{u}^T \mathbf{W}_\ell \mathbf{v} = \langle \mathbf{V}_\ell^T \mathbf{u}, \mathbf{V}_\ell^T \mathbf{v} \rangle_{\mathbf{I}_{n_\ell}}, \quad (3.13)$$

où  $\langle \cdot, \cdot \rangle_{\mathbf{I}_{n_\ell}}$  désigne le produit scalaire canonique de  $\mathbb{R}^{n_\ell}$ . On note  $\| \cdot \|_{\mathbf{W}_\ell}$  la norme induite par le produit scalaire  $\langle \cdot, \cdot \rangle_{\mathbf{W}_\ell}$ . Les espaces de Hilbert  $H_\ell$  d'entrée et de sortie des simulateurs  $f_\ell$  des équations (3.1) et (3.2) sont donc  $H_\ell = (\mathbb{R}^{n_\ell}, \langle \cdot, \cdot \rangle_{\mathbf{W}_\ell})$ . On note  $\mathbf{X}_\ell$  (resp.  $\mathbf{Y}_\ell$ ) le vecteur aléatoire de moment d'ordre 2 fini d'entrée (resp. de sortie) du modèle  $f_\ell$ . On cherche toujours à estimer  $\boldsymbol{\mu}_L := \mathbb{E}[\mathbf{Y}_L]$ , l'espérance de la sortie du modèle haute fidélité. Ce contexte est un cas particulier de la section 3.1 où  $H_\ell = (\mathbb{R}^{n_\ell}, \langle \cdot, \cdot \rangle_{\mathbf{W}_\ell})$ . Les différentes expressions de la section 3.1.2 restent donc inchangées, on notera simplement  $\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}$  l'estimateur MLMC de l'équation (3.7) pour des champs discrétisés représentés par des vecteurs.

### 3.2.2 Grilles centrées sur les cellules et opérateurs de transfert linéaires

Le choix des opérateurs de transfert utilisés dépend des grilles considérées et des représentations des champs sur ces grilles. Afin de faciliter l'étude, on se place sur un domaine unidimensionnel  $\mathcal{D} := [0, 1]$ . On considère  $L + 1$  différents maillages de  $\mathcal{D}$  avec des grilles uniformes composées de  $n_\ell$  cellules, chaque cellule étant alors de taille  $n_\ell^{-1}$ . Les matrices de Gram associées à ces grilles sont simplement définies comme  $\mathbf{W}_\ell = n_\ell^{-1} \mathbf{I}_{n_\ell}$ , et on a  $\mathbf{V}_\ell = n_\ell^{-1/2} \mathbf{I}_{n_\ell}$ . On fixe, de plus, la relation  $n_\ell = 2n_{\ell-1}$  pour  $\ell = 1, \dots, L$  entre les différentes grilles. La représentation d'un champ sur une grille composée de  $n_\ell$  cellules est un vecteur de  $\mathbb{R}^{n_\ell}$  dont les éléments correspondent à la valeur du champ au centre des différentes cellules. Un exemple de hiérarchie de 3 grilles est montré dans la figure 3.1.

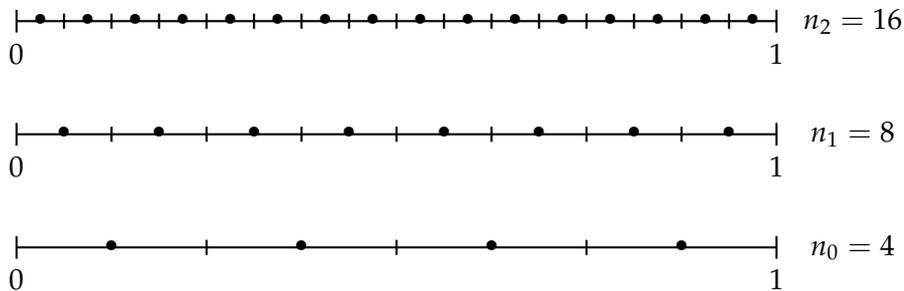


FIGURE 3.1 – Exemple de 3 grilles de taille 4, 8 et 16. Les points représentent le centre des cellules.

Il est important de remarquer que, par construction, les grilles ne sont pas imbriquées, c'est-à-dire que les points de l'espace où l'on connaît la valeur du champ ne sont pas les mêmes d'une grille à l'autre. Cette propriété découle du choix d'utiliser

des grilles centrées sur les cellules avec un facteur de réduction de 2 entre les grilles. Un tel choix est fait car c'est le contexte qui se rapproche le plus de celui considéré pour l'application présentée dans le chapitre 5. Dans les méthodes multigrilles, il est connu que l'utilisation de grilles centrées sur les cellules est un contexte plus complexe que les grilles centrées sur les nœuds et a fait l'objet de plusieurs travaux [33, 36, 51, 74].

On cherche maintenant à définir des opérateurs de prolongation et de restriction pour passer d'une grille à l'autre. On choisit de considérer des opérateurs de transfert linéaires pour leur simplicité et le faible coût de calcul de leur application (on rappelle que l'on a supposé que leur coût était négligeable comparé à celui de l'application du modèle). On identifie donc l'opérateur de prolongation  $P_\ell^{\ell'}$  (où  $n_{\ell'} > n_\ell$ ) avec la matrice  $\mathbf{P}_\ell^{\ell'} \in \mathbb{R}^{n_{\ell'} \times n_\ell}$  telle que  $P_\ell^{\ell'}(\mathbf{x}_\ell) = \mathbf{P}_\ell^{\ell'} \mathbf{x}_\ell$  pour tout  $\mathbf{x}_\ell \in \mathbb{R}^{n_\ell}$ . De la même manière, on identifie  $R_\ell^{\ell'}$  avec la matrice  $\mathbf{R}_\ell^{\ell'} \in \mathbb{R}^{n_\ell \times n_{\ell'}}$ . Le choix des opérateurs de transfert va être guidé par la vérification de propriétés essentielles. On cherche, pour la prolongation, à conserver la norme d'un vecteur à travers différentes grilles, i.e., pour  $n_{\ell'} > n_\ell$ ,

$$\|\mathbf{P}_\ell^{\ell'} \mathbf{x}_\ell\|_{\mathbf{W}_{\ell'}} = \|\mathbf{x}_\ell\|_{\mathbf{W}_\ell} \quad \forall \mathbf{x}_\ell \in \mathbb{R}^{n_\ell}. \quad (3.14)$$

Cette propriété peut être réécrite sous la forme

$$\mathbf{x}_\ell^\top (\mathbf{P}_\ell^{\ell'})^\top \mathbf{W}_{\ell'} \mathbf{P}_\ell^{\ell'} \mathbf{x}_\ell = \mathbf{x}_\ell^\top \mathbf{W}_\ell \mathbf{x}_\ell, \quad \forall \mathbf{x}_\ell \in \mathbb{R}^{n_\ell}, \quad (3.15)$$

et se simplifie donc en  $(\mathbf{P}_\ell^{\ell'})^\top \mathbf{W}_{\ell'} \mathbf{P}_\ell^{\ell'} = \mathbf{W}_\ell$ . Étant donné que  $\mathbf{W}_\ell = n_\ell^{-1} \mathbf{I}_{n_\ell}$  pour  $\ell = 0, \dots, L$ , cela revient à imposer  $(\mathbf{P}_\ell^{\ell'})^\top \mathbf{P}_\ell^{\ell'} = (n_{\ell'}/n_\ell) \mathbf{I}_{n_\ell}$  avec, ici,  $n_{\ell'}/n_\ell = 2^{\ell'-\ell}$ . Pour l'opérateur de restriction on impose la relation

$$\mathbf{R}_\ell^{\ell'} = \alpha^{\ell'-\ell} (\mathbf{P}_\ell^{\ell'})^\top, \quad \alpha \in \mathbb{R}. \quad (3.16)$$

Cette relation est similaire à la propriété variationnelle imposée dans les méthodes multigrilles [7, Chapitre 3]. Le choix du facteur  $\alpha$  n'est pas trivial et sera discuté par la suite. Ce que l'on impose ici pour contraindre ce facteur  $\alpha$ , et qui découle du problème étudié dans ce manuscrit, est que la restriction d'un vecteur aléatoire suivant une loi normale standard multivariée suit également une loi normale standard multivariée. Autrement dit, on impose

$$\mathbf{R}_\ell^{\ell'} \mathbf{X}_{\ell'} \sim \mathcal{N}(\mathbf{0}_{n_\ell}, \mathbf{I}_{n_\ell}), \quad \forall \mathbf{X}_{\ell'} \sim \mathcal{N}(\mathbf{0}_{n_{\ell'}}, \mathbf{I}_{n_{\ell'}}). \quad (3.17)$$

Étant donné que l'opérateur de restriction est linéaire, on a bien  $\mathbb{E}[\mathbf{R}_\ell^{\ell'} \mathbf{X}_{\ell'}] = \mathbf{0}_{n_\ell}$ . La covariance de  $\mathbf{R}_\ell^{\ell'} \mathbf{X}_{\ell'}$  s'écrit

$$\mathbf{C}[\mathbf{R}_\ell^{\ell'} \mathbf{X}_{\ell'}] = \mathbf{R}_\ell^{\ell'} \mathbf{X}_{\ell'} (\mathbf{R}_\ell^{\ell'} \mathbf{X}_{\ell'})^\top = \mathbf{R}_\ell^{\ell'} \mathbf{X}_{\ell'} \mathbf{X}_{\ell'}^\top (\mathbf{R}_\ell^{\ell'})^\top = \mathbf{R}_\ell^{\ell'} (\mathbf{R}_\ell^{\ell'})^\top, \quad (3.18)$$

il suffit donc de vérifier  $\mathbf{R}_\ell^{\ell'} (\mathbf{R}_\ell^{\ell'})^\top = \mathbf{I}_{n_\ell}$  pour obtenir une variance unitaire.

Pour respecter ces différentes propriétés on choisit spécifiquement

$$\mathbf{P}_{\ell-1}^\ell := \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \\ & & & & \ddots \end{bmatrix} \in \mathbb{R}^{n_\ell \times n_{\ell-1}}, \quad (3.19)$$

et

$$\mathbf{R}_\ell^{\ell-1} := \frac{1}{\sqrt{2}}(\mathbf{P}_{\ell-1}^\ell)^\top = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & & \ddots & \\ & & & & \ddots \end{bmatrix} \in \mathbb{R}^{n_{\ell-1} \times n_\ell}. \quad (3.20)$$

On a alors bien  $(\mathbf{P}_{\ell-1}^\ell)^\top \mathbf{P}_{\ell-1}^\ell = 2\mathbf{I}_{n_{\ell-1}}$  et  $\mathbf{R}_\ell^{\ell-1}(\mathbf{R}_\ell^{\ell-1})^\top = \mathbf{I}_{n_{\ell-1}}$  pour  $\ell = 1, \dots, L$ . La figure 3.2 montre un exemple d'application des opérateurs de transfert sur un vecteur  $\mathbf{x}$ .

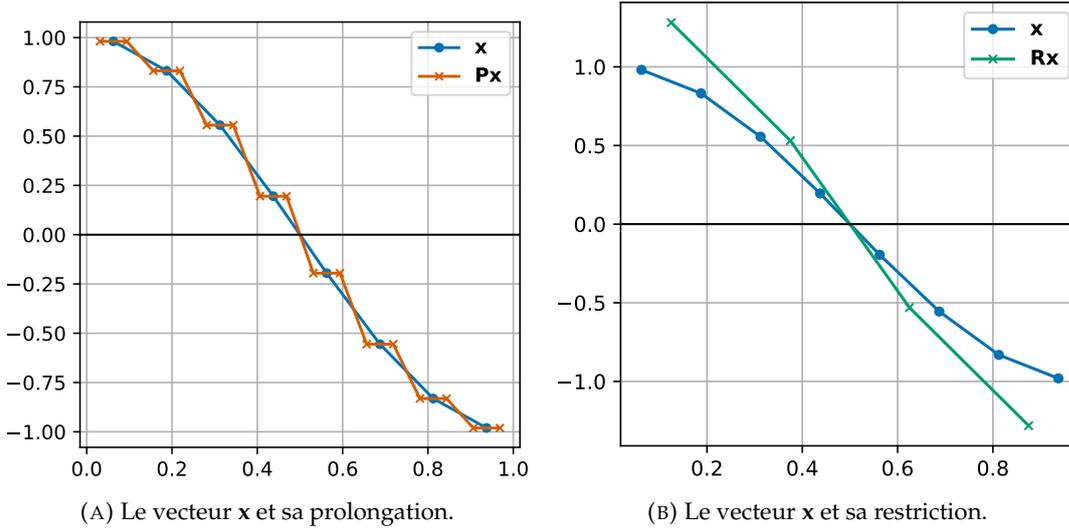


FIGURE 3.2 – Affichage du vecteur  $\mathbf{x}$  (bleu) défini comme la représentation discrétisé de  $f(x) = \cos(\pi x)$  sur une grille de taille 8 centrée sur les cellules, ainsi que sa prolongation  $\mathbf{P}\mathbf{x} \in \mathbb{R}^{16}$  (orange) et sa restriction  $\mathbf{R}\mathbf{x} \in \mathbb{R}^4$  (vert).

En utilisant la définition des opérateurs de prolongation et de restriction entre un espace  $H_\ell$  et l'espace  $H_L$  (cf. équations (3.4) et (3.5)), on peut construire les matrices  $\mathbf{P}_\ell^L \in \mathbb{R}^{n_L \times n_\ell}$  et  $\mathbf{R}_\ell^L \in \mathbb{R}^{n_\ell \times n_L}$ , pour  $\ell = 0, \dots, L-1$ . On pose  $\mathbf{P}_L^L = \mathbf{R}_L^L = \mathbf{I}_{n_L}$  la matrice identité. Avec ces opérateurs fixés, il devient possible, pour une hiérarchie de modèles  $(f_\ell)_{\ell=0}^L$  donnée, de spécifier la hiérarchie  $(\tilde{f}_\ell)_{\ell=0}^L$  définie dans l'équation (3.6) et d'estimer  $\mu_L$  à l'aide de l'estimateur MLMC (3.7).

### 3.3 Analyse spectrale

L'utilisation de grilles de différentes fidélités et d'opérateurs de transfert entre ces grilles est à la base des méthodes multigrilles [69, 7, 76] et, dans un tel contexte, l'impact de ces opérateurs a été fortement étudié. Les effets de ces opérateurs sont notamment étudiés sous le prisme d'une analyse spectrale, c'est à dire en décomposant le champ discrétisé considéré en composantes associées aux différentes échelles spatiales représentables. Nous adaptons cet outil à l'estimateur MLMC (3.7) afin d'étudier les effets des opérateurs de transfert (3.19) et (3.20) et leur impact sur la variance de l'estimateur.

#### 3.3.1 Base de Hartley

On commence par s'intéresser à la décomposition d'un champ discrétisé en différentes composantes associées aux échelles spatiales représentables. La décomposition

discrète de Fourier est courante pour une telle étude, mais il est possible d'utiliser d'autres décompositions, par exemple une base de sinus dans [7]. On propose d'utiliser la base de Hartley [5, 6], définie comme la somme des parties réelles et imaginaires de la base de Fourier et qui a pour avantage que les vecteurs de la base sont tous réels. L'absence de complexes facilite la visualisation des différents vecteurs de la base et l'interprétation des différentes échelles. On adapte la base de Hartley de [5] pour des grilles centrées sur les cellules en définissant les matrices de Hartley  $\mathbf{H}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$  pour  $\ell = 0, \dots, L$  par

$$(\mathbf{H}_\ell)_{j,k} := \frac{1}{\sqrt{n_\ell}} \left( \cos \frac{2(j + \frac{1}{2})k\pi}{n_\ell} + \sin \frac{2(j + \frac{1}{2})k\pi}{n_\ell} \right), \quad \forall j, k = 0, \dots, n_\ell - 1. \quad (3.21)$$

Les matrices  $\mathbf{H}_\ell$  sont orthogonales par rapport au produit scalaire canonique pour tout  $\ell = 0, \dots, L$ , i.e.,  $\mathbf{H}_\ell^T \mathbf{H}_\ell = \mathbf{H}_\ell \mathbf{H}_\ell^T = \mathbf{I}_{n_\ell}$  (cf. section A.1). Il faut, cependant, noter qu'elles ne sont pas symétriques. On désigne par  $\mathbf{h}_k^\ell$  la colonne  $k$  de la matrice  $\mathbf{H}_\ell$  pour  $k = 0, \dots, n_\ell - 1$ . Les colonnes de cette matrice forment la base de Hartley centrée sur les cellules. La figure 3.3 présente les colonnes de la matrice de Hartley  $\mathbf{H}_0$  de taille  $n_0 = 16$  et permet de visualiser les différentes échelles représentables sur une grille de taille  $n_0$ .

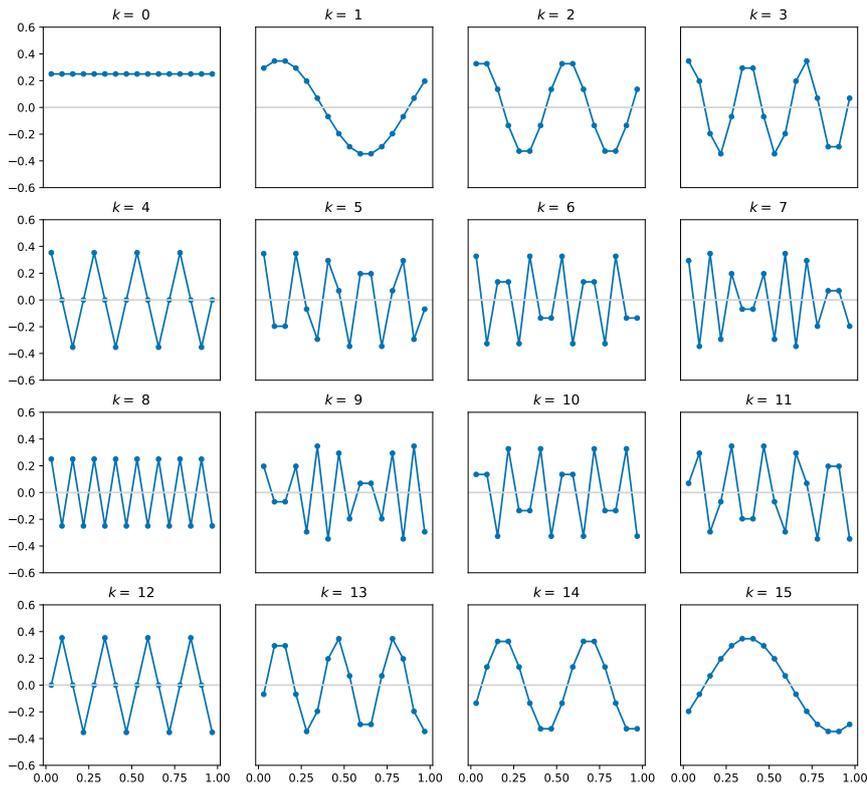


FIGURE 3.3 – Les colonnes  $\mathbf{h}_k^0$  de la matrice de Hartley  $\mathbf{H}_0$  où  $n_0 = 16$ . Elles sont représentées sur une grille qui discrétise  $[0, 1]$  en  $n_0$  cellules. Chaque colonne correspond à une échelle représentable sur la grille associée.

On appelle grandes échelles (ou basses fréquences) les composantes associées aux colonnes  $\mathbf{h}_k^\ell$  avec  $k < n_\ell/4$  ou  $k \geq 3n_\ell/4$ . Les composantes associées aux colonnes  $\mathbf{h}_k^\ell$  avec  $3n_\ell/4 > k \geq n_\ell/4$  sont les fines, ou petites échelles (ou hautes fréquences). Par définition d'une base, tout vecteur de  $\mathbb{R}^{n_\ell}$  peut s'écrire comme une combinaison

linéaire de colonnes de  $\mathbf{H}_\ell$  :

$$\forall \mathbf{x}_\ell \in \mathbb{R}^{n_\ell}, \quad \exists \alpha_0, \dots, \alpha_{n_\ell-1} \in \mathbb{R}; \quad \mathbf{x}_\ell = \sum_{k=0}^{n_\ell-1} \alpha_k \mathbf{h}_k^\ell. \quad (3.22)$$

Pour certaines figures ou représentations il est plus pertinent de réordonner les colonnes de  $\mathbf{H}_\ell$  par échelle représentable de longueur décroissante, i.e., par fréquence croissante. On définit la matrice de Hartley réordonnée

$$\mathbf{H}_\ell^* := \begin{bmatrix} \mathbf{h}_0^\ell & \mathbf{h}_{n_\ell-1}^\ell & \mathbf{h}_1^\ell & \mathbf{h}_{n_\ell-2}^\ell & \cdots & \mathbf{h}_{n_\ell/2-1}^\ell & \mathbf{h}_{n_\ell/2}^\ell \end{bmatrix}. \quad (3.23)$$

Cette matrice  $\mathbf{H}_\ell^*$  est obtenue en multipliant  $\mathbf{H}_\ell$  à droite par la matrice de permutation  $\mathbf{\Pi} = (\Pi_{j,k})_{j,k=0}^{n_\ell-1}$  avec  $\Pi_{j,2k} = \delta_{j,k}$  et  $\Pi_{j,2k+1} = \delta_{j,n_\ell-k-1}$  pour  $j = 0, \dots, n_\ell - 1$  et  $k = 0, \dots, n_\ell/2 - 1$ . Les vecteurs réordonnés sont présentés sur la figure 3.4.

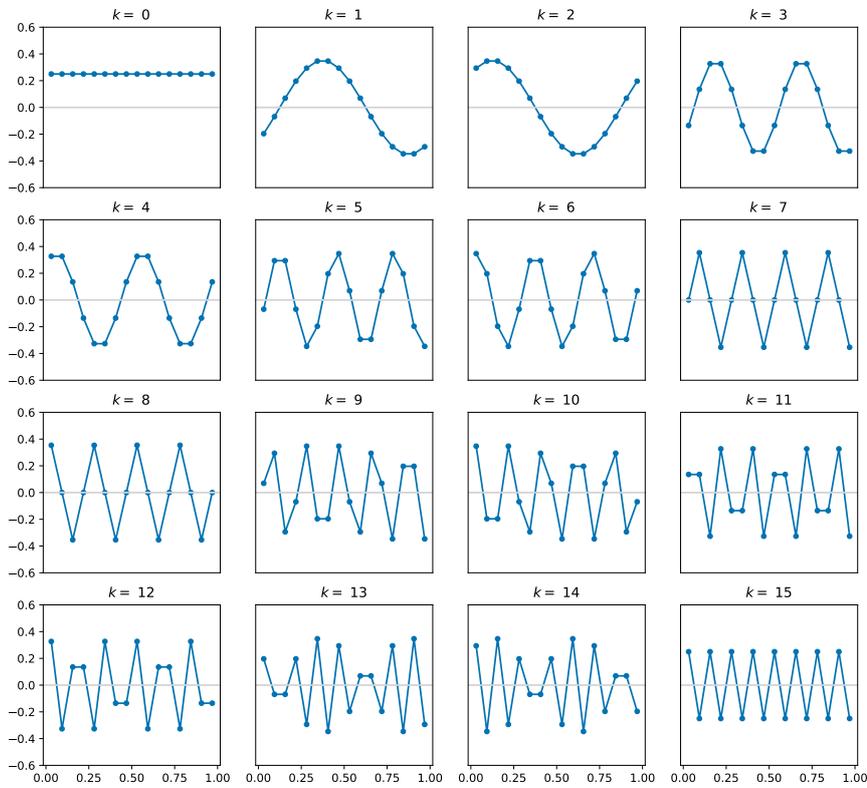


FIGURE 3.4 – Les colonnes  $(\mathbf{h}^*)_k^0$  de la matrice de Hartley réordonnée  $\mathbf{H}_0^*$  où  $n_0 = 16$ . Elles sont représentées sur une grille qui discrétise  $[0, 1]$  en  $n_0$  cellules. Les fréquences représentées sont, ici, croissantes avec  $k$ .

La base de Hartley permet ainsi de décomposer un vecteur en une somme des contributions des différentes échelles. En exploitant son orthogonalité, on peut utiliser  $\mathbf{H}_L$  pour décomposer la MSE de l'estimateur MLMC  $\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}$ , donnée dans l'équation (3.8),

$$\text{MSE}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}, \boldsymbol{\mu}) = \mathbb{E}[\|\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu}\|_{\mathbf{W}_L}^2] \quad (3.24)$$

$$= \mathbb{E}[\|\mathbf{V}_L^T(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})\|_{\mathbf{I}_{n_L}}^2] \quad (3.25)$$

$$= \mathbb{E}[\|\mathbf{H}_L^T \mathbf{V}_L^T(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})\|_{\mathbf{I}_{n_L}}^2] \quad (3.26)$$

En utilisant le fait que, sur les grilles régulières, on a  $\mathbf{W}_L = n_L^{-1} \mathbf{I}_{n_L}$ , on peut exprimer la MSE de la manière suivante :

$$\text{MSE}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}, \boldsymbol{\mu}) = \mathbb{E}[\|\mathbf{H}_L^T \mathbf{V}_L^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})\|_{\mathbf{I}_{n_L}}^2] \quad (3.27)$$

$$= \mathbb{E}\left[\sum_{k=0}^{n_L-1} (\mathbf{H}_L^T \mathbf{V}_L^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu}))_k^2\right] \quad (3.28)$$

$$= \mathbb{E}\left[\sum_{k=0}^{n_L-1} \frac{1}{n_L} \left(\sum_{j=0}^{n_L-1} (\mathbf{H}_L)_{k,j}^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})_j\right)^2\right] \quad (3.29)$$

$$= \mathbb{E}\left[\sum_{k=0}^{n_L-1} \frac{1}{n_L} \left(\sum_{j=0}^{n_L-1} (\mathbf{H}_L)_{j,k} (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})_j\right)^2\right] \quad (3.30)$$

$$= \mathbb{E}\left[\sum_{k=0}^{n_L-1} \frac{1}{n_L} \left((\mathbf{h}_k^L)^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})\right)^2\right] \quad (3.31)$$

$$= \sum_{k=0}^{n_L-1} \frac{1}{n_L} \mathbb{E}\left[\left((\mathbf{h}_k^L)^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})\right)^2\right] \quad (3.32)$$

$$= \sum_{k=0}^{n_L-1} v_k \quad (3.33)$$

$$= \|\mathbf{v}\|_1, \quad (3.34)$$

avec  $\mathbf{v} = (v_k)_{k=0}^{n_L-1}$  le vecteur de la MSE spectrale donc les éléments sont  $v_k := n_L^{-1} \mathbb{E}\left[\left((\mathbf{h}_k^L)^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu})\right)^2\right]$ . Chacun des éléments de ce vecteur correspond à la MSE d'une des composante d'échelle de l'estimateur dans la base  $\mathbf{H}_L$  et la MSE totale est la somme de ces coefficients. On rappelle que l'on s'intéresse à la MSE par rapport à  $\boldsymbol{\mu}_L = \mathbb{E}[\mathbf{X}_L]$  et on a  $\mathbb{E}[\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}] = \boldsymbol{\mu}_L$ . On a alors

$$v_k = \frac{1}{n_L} \mathbb{E}\left[\left((\mathbf{h}_k^L)^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \boldsymbol{\mu}_L)\right)^2\right] \quad (3.35)$$

$$= \frac{1}{n_L} \mathbb{E}\left[\left((\mathbf{h}_k^L)^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \mathbb{E}[\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}])\right)^2\right] \quad (3.36)$$

$$= \frac{1}{n_L} \mathbb{E}\left[\left((\mathbf{h}_k^L)^T (\hat{\boldsymbol{\mu}}_L^{\text{MLMC}} - \mathbb{E}[(\mathbf{h}_k^L)^T \hat{\boldsymbol{\mu}}_L^{\text{MLMC}}])\right)^2\right] \quad (3.37)$$

$$= \frac{1}{n_L} \mathbf{V}[(\mathbf{h}_k^L)^T \hat{\boldsymbol{\mu}}_L^{\text{MLMC}}]. \quad (3.38)$$

La MSE est alors égale à la variance de l'estimateur

$$\text{MSE}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}, \boldsymbol{\mu}_L) = \mathcal{V}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}) = \|\mathbf{v}\|_1 = \sum_{k=0}^{n_L-1} v_k, \quad (3.39)$$

avec  $v_k = n_L^{-1} \mathbf{V}[(\mathbf{h}_k^L)^T \hat{\boldsymbol{\mu}}_L^{\text{MLMC}}]$  et le vecteur  $\mathbf{v}$  que l'on appelle vecteur de variance spectrale.

### 3.3.2 Effets des opérateurs de transfert sur la base de Hartley

On souhaite étudier les effets de l'application des opérateurs de transfert sur les différentes composantes d'échelle d'un champ discrétisé. Ces effets ont déjà été le sujet de nombreux travaux pour les méthodes multigrilles [7, 69, 76]. On étudie ici

l'effet des opérateurs de transfert spécifiques des équations (3.19) et (3.20) sur les vecteurs de la base de Hartley (3.21).

### Prolongation de la base de Hartley

On présente les résultats pour un niveau  $\ell$  fixé entre 1 et  $L$  et pour l'opérateur de prolongation  $\mathbf{P} := \mathbf{P}_{\ell-1}^\ell$ . Pour  $j, k = 0, \dots, n_{\ell-1} - 1$ ,

$$(\mathbf{P}\mathbf{h}_k^{\ell-1})_{2j} = (\mathbf{P}\mathbf{h}_k^{\ell-1})_{2j+1} = (\mathbf{h}_k^{\ell-1})_j \quad (3.40)$$

$$= \frac{1}{\sqrt{n_{\ell-1}}} \left( \cos \frac{2(j + \frac{1}{2})k\pi}{n_{\ell-1}} + \sin \frac{2(j + \frac{1}{2})k\pi}{n_{\ell-1}} \right) \quad (3.41)$$

$$= \frac{\sqrt{2}}{\sqrt{n_\ell}} \left( \cos \frac{(4j + 2)k\pi}{n_\ell} + \sin \frac{(4j + 2)k\pi}{n_\ell} \right), \quad (3.42)$$

car  $n_\ell = 2n_{\ell-1}$ . En décomposant  $(4j + 2)k\pi = (4j + 1)k\pi + k\pi$  et en utilisant les relations trigonométriques sur le cosinus ou le sinus d'une somme, on obtient,

$$(\mathbf{P}\mathbf{h}_k^{\ell-1})_{2j} = \sqrt{2} \cos \frac{k\pi}{n_\ell} (\mathbf{h}_k^\ell)_{2j} + \frac{\sqrt{2}}{\sqrt{n_\ell}} \left[ \sin \frac{k\pi}{n_\ell} \left( \cos \frac{(4j + 1)k\pi}{n_\ell} - \sin \frac{(4j + 1)k\pi}{n_\ell} \right) \right]. \quad (3.43)$$

On note  $c_k := \cos(k\pi/n_\ell)$  dans la suite. Le second terme de cette expression peut être simplifié en utilisant les relations suivantes

$$\sin \frac{k\pi}{n_\ell} = \sin \left( \frac{(n_{\ell-1} + k)\pi}{2n_{\ell-1}} - \frac{\pi}{2} \right) = -c_{n_{\ell-1}+k}, \quad (3.44)$$

$$\cos \frac{(4j + 1)k\pi}{n_\ell} = \cos \left( \frac{(4j + 1)(n_{\ell-1} + k)\pi}{n_\ell} - \frac{\pi}{2} \right) = \sin \frac{(4j + 1)(n_{\ell-1} + k)\pi}{n_\ell}, \quad (3.45)$$

$$\sin \frac{(4j + 1)k\pi}{n_\ell} = \sin \left( \frac{(4j + 1)(n_{\ell-1} + k)\pi}{n_\ell} - \frac{\pi}{2} \right) = -\cos \frac{(4j + 1)(n_{\ell-1} + k)\pi}{n_\ell}. \quad (3.46)$$

On obtient alors que

$$\frac{\sqrt{2}}{\sqrt{n_\ell}} \left[ \sin \frac{k\pi}{n_\ell} \left( \cos \frac{(4j + 1)k\pi}{n_\ell} - \sin \frac{(4j + 1)k\pi}{n_\ell} \right) \right] = -\sqrt{2}c_{n_{\ell-1}+k}(\mathbf{h}_{n_{\ell-1}+k}^\ell)_{2j}, \quad (3.47)$$

et on peut conclure que  $(\mathbf{P}\mathbf{h}_k^{\ell-1})_{2j} = \sqrt{2}(c_k(\mathbf{h}_k^\ell)_{2j} - c_{n_{\ell-1}+k}(\mathbf{h}_{n_{\ell-1}+k}^\ell)_{2j})$ . De la même manière on a  $(\mathbf{P}\mathbf{h}_k^{\ell-1})_{2j+1} = \sqrt{2}(c_k(\mathbf{h}_k^\ell)_{2j+1} - c_{n_{\ell-1}+k}(\mathbf{h}_{n_{\ell-1}+k}^\ell)_{2j+1})$ . Ces résultats peuvent être réécrit vectoriellement,

$$\mathbf{P}\mathbf{h}_k^{\ell-1} = \sqrt{2}(c_k\mathbf{h}_k^\ell - c_{n_{\ell-1}+k}\mathbf{h}_{n_{\ell-1}+k}^\ell), \quad \forall k = 0, \dots, n_{\ell-1} - 1. \quad (3.48)$$

Une visualisation de ce résultat est présenté sur la figure 3.5 avec la prolongation des vecteurs de la base de Hartley de taille 8.

La prolongation d'un vecteur colonne  $\mathbf{h}_k^{\ell-1}$  de la base associée à une grille grossière aboutit à une combinaison linéaire de deux vecteurs colonnes de la base sur la grille fine,  $\mathbf{h}_k^\ell$  et  $\mathbf{h}_{n_{\ell-1}+k}^\ell$  pour  $k = 0, \dots, n_{\ell-1}$ . Les coefficients intervenants dans cette

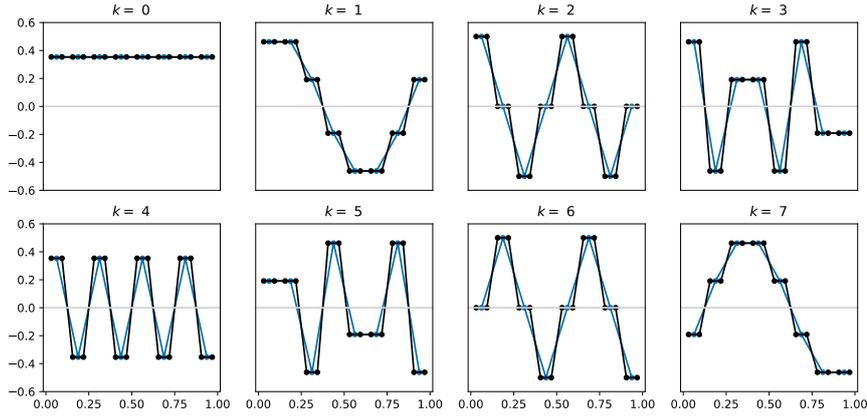


FIGURE 3.5 – Les colonnes  $\mathbf{h}_k^0$  (bleu) de la matrice de Hartley  $\mathbf{H}_0$  de taille  $n_0 = 8$  et leur prolongation  $\mathbf{P}\mathbf{h}_k^0$  (noir) qui sont donc des vecteurs de taille  $n_1 = 16$ .

combinaison linéaire sont strictement décroissants selon  $k$ , i.e.,

$$1 = c_0 > c_1 > \cdots > (c_{n_{\ell-1}} = 0) > \cdots > c_{2n_{\ell-1}-1} > -1. \quad (3.49)$$

Dans cette combinaison linéaire, l'un des deux vecteurs,  $\mathbf{h}_k^\ell$  ou  $\mathbf{h}_{n_{\ell-1}+k}^\ell$  représente la même échelle que le vecteur prolongé  $\mathbf{h}_k^{\ell-1}$ . En effet, pour  $k < n_{\ell-1}/2$ , les vecteurs  $\mathbf{h}_k^{\ell-1}$  et  $\mathbf{h}_k^\ell$  correspondent à la même échelle, tandis que pour  $k \geq n_{\ell-1}/2$ , c'est le vecteur  $\mathbf{h}_{n_{\ell-1}+k}^\ell$  qui correspond à la même échelle que  $\mathbf{h}_k^{\ell-1}$ . Cela peut se constater avec les figures 3.3 et 3.5 où, par exemple, la prolongation de  $\mathbf{h}_1^0$  est une combinaison linéaire de  $\mathbf{h}_1^1$  et  $\mathbf{h}_9^1$  avec  $\mathbf{h}_1^1$  qui représente la même échelle que  $\mathbf{h}_1^0$ . L'autre colonne correspond alors toujours à une échelle plus petite (ou une fréquence plus élevée) et ajoute donc une échelle parasite dans le résultat de la prolongation. Heureusement, ce terme est celui avec la plus faible amplitude. En effet,  $c_k$  tend vers 1 et  $c_{n_{\ell-1}+k}$  vers 0 lorsque  $k$  tend vers 0. À l'inverse,  $-c_{n_{\ell-1}+k}$  tend vers 1 et  $c_k$  vers 0 lorsque  $k$  tend vers  $n_{\ell-1}$ . La prolongation d'un vecteur  $\mathbf{h}_k^{\ell-1}$  introduit donc un terme parasite associé à une petite échelle qui n'est pas représentable sur la grille grossière de taille  $n_{\ell-1}$ . Ce terme parasite est d'amplitude plus faible que celui associé aux grandes échelles mais constitue, néanmoins, une erreur dans le vecteur prolongé.

### Restriction de la base de Hartley

Pour la restriction  $\mathbf{R} := \mathbf{R}_\ell^{\ell-1}$  et pour  $j = 0 \dots, n_{\ell-1} - 1, k = 0, \dots, 2n_{\ell-1} - 1$ ,

$$(\mathbf{R}\mathbf{h}_k^\ell)_j = \frac{1}{\sqrt{2}}((\mathbf{h}_k^\ell)_{2j} + (\mathbf{h}_k^\ell)_{2j+1}) \quad (3.50)$$

$$= \frac{1}{\sqrt{2n_\ell}} \left[ \cos \frac{(4j+1)k\pi}{n_\ell} + \cos \frac{(4j+3)k\pi}{n_\ell} + \sin \frac{(4j+1)k\pi}{n_\ell} + \sin \frac{(4j+3)k\pi}{n_\ell} \right] \quad (3.51)$$

$$= \frac{1}{\sqrt{2n_\ell}} \left[ 2 \cos \frac{(8j+4)k\pi}{2n_\ell} \cos \frac{2k\pi}{2n_\ell} + 2 \sin \frac{(8j+4)k\pi}{2n_\ell} \cos \frac{2k\pi}{2n_\ell} \right] \quad (3.52)$$

$$= \frac{c_k}{\sqrt{n_{\ell-1}}} \left[ \cos \frac{(2j+1)k\pi}{n_{\ell-1}} + \sin \frac{(2j+1)k\pi}{n_{\ell-1}} \right]. \quad (3.53)$$

Pour  $k = 0, \dots, n_{\ell-1} - 1$  on a alors directement  $\mathbf{R}\mathbf{h}_k^\ell = c_k \mathbf{h}_k^{\ell-1}$ . Pour les autres colonnes de la matrice de Hartley on a, pour  $k = 0, \dots, n_{\ell-1} - 1$ ,

$$(\mathbf{R}\mathbf{h}_{n_{\ell-1}+k}^\ell)_j = \frac{c_{n_{\ell-1}+k}}{\sqrt{n_{\ell-1}}} \left[ \cos\left(\frac{(2j+1)k\pi}{n_{\ell-1}} + \pi\right) + \sin\left(\frac{(2j+1)k\pi}{n_{\ell-1}} + \pi\right) \right] \quad (3.54)$$

$$= -c_{n_{\ell-1}+k} (\mathbf{h}_k^{\ell-1})_j. \quad (3.55)$$

Ce qui, vectoriellement, donne

$$\mathbf{R}\mathbf{h}_k^\ell = c_k \mathbf{h}_k^{\ell-1} \quad \text{et} \quad \mathbf{R}\mathbf{h}_{n_{\ell-1}+k}^\ell = -c_{n_{\ell-1}+k} \mathbf{h}_k^{\ell-1}, \quad \forall k = 0, \dots, n_{\ell-1} - 1. \quad (3.56)$$

La figure 3.6 présente une visualisation de ce résultat avec la restriction des vecteurs de la base de Hartley de taille 16.

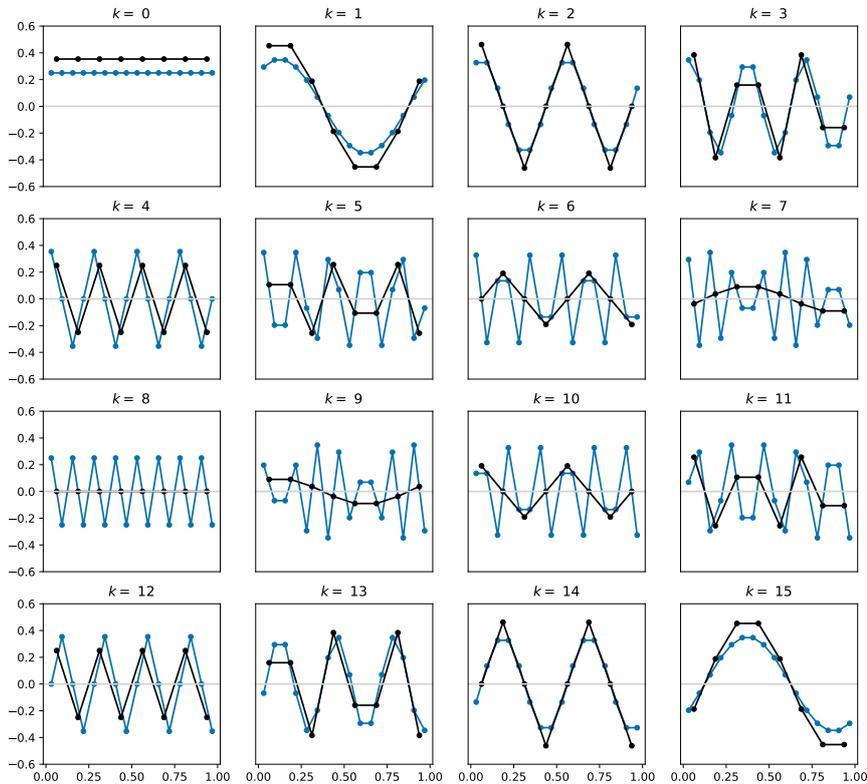


FIGURE 3.6 – Les colonnes  $\mathbf{h}_k^1$  (bleu) de la matrice de Hartley  $\mathbf{H}_1$  de taille  $n_1 = 16$  et leur restriction  $\mathbf{R}\mathbf{h}_k^1$  (noir) qui sont donc des vecteurs de taille  $n_0 = 8$ .

Pour  $k < n_{\ell-1}$ , la restriction d'un vecteur  $\mathbf{h}_k^\ell$  de la base  $\mathbf{H}_\ell$  sur la grille fine correspond au  $k$ -ième vecteur de la base  $\mathbf{H}_{\ell-1}$  sur la grille grossière, réduit par le facteur  $c_k$  qui est compris entre 0 et 1. Ce vecteur  $\mathbf{h}_k^{\ell-1}$  obtenu représente la même échelle si  $k < n_{\ell-1}/2$ , et représente une échelle plus grande si  $k \geq n_{\ell-1}/2$ . Par exemple, en comparant les figures 3.3 et 3.6, on a bien les restrictions de  $\mathbf{h}_0^1, \mathbf{h}_1^1, \mathbf{h}_2^1$  et  $\mathbf{h}_3^1$  qui représentent les mêmes échelles, tandis que pour  $\mathbf{h}_4^1, \mathbf{h}_5^1, \mathbf{h}_6^1$  et  $\mathbf{h}_7^1$ , dont les échelles ne sont pas représentables sur la grille de taille  $n_0$ , leur restriction correspond à des échelles plus grandes. Pour  $k \geq n_{\ell-1}$  les résultats sont similaires. La restriction du vecteur  $\mathbf{h}_{n_{\ell-1}+k}^\ell$  est le vecteur complémentaire  $\mathbf{h}_k^{\ell-1}$  de la grille grossière réduit par le facteur  $-c_{n_{\ell-1}+k}$  compris entre 0 et 1. Pour  $k \geq 3n_{\ell-1}/2$  le vecteur obtenu représente bien la même échelle et pour  $k < 3n_{\ell-1}/2$  une échelle plus grande. Heureusement,



### 3.3.3 Impact sur la variance du MLMC

En ayant une meilleure idée des effets des opérateurs de transfert de grille, on cherche maintenant à connaître les conséquences de ces effets sur l'estimateur MLMC, et plus particulièrement sur sa MSE. On rappelle que, dans notre cas, la MSE est seulement constituée de la variance de l'estimateur car on estime l'espérance de la sortie du modèle haute fidélité. Pour étudier plus en détail la variance de l'estimateur MLMC présenté dans l'équation (3.7) avec les modèles  $\tilde{f}_\ell$  définis par l'équation (3.6), nous allons formuler des hypothèses sur les modèles  $f_\ell$ .

#### Modèles linéaires

On suppose ici que les applications  $f_\ell$  sont linéaires pour  $\ell = 0, \dots, L$  et que l'on peut donc les identifier avec des matrices  $\mathbf{F}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ . Par conséquent, les modèles  $\tilde{f}_\ell$  sont également linéaires et peuvent être identifiés avec les matrices

$$\tilde{\mathbf{F}}_\ell := \mathbf{P}_\ell^L \mathbf{F}_\ell \mathbf{R}_L^\ell \in \mathbb{R}^{n_L \times n_L} \quad (3.63)$$

pour  $\ell = 0, \dots, L$  et avec  $\mathbf{P}_L^L = \mathbf{R}_L^L = \mathbf{I}_{n_L}$ . Ces matrices correspondent aux simulateurs utilisés dans l'estimateur MLMC (3.7) de  $\mathbb{E}[\mathbf{F}_L \mathbf{X}_L]$  avec  $\mathbf{X}_L$  un vecteur aléatoire. On a alors, en partant de l'équation (3.11),

$$\mathcal{V}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}) = \frac{1}{M_0} \mathcal{V}(\tilde{\mathbf{F}}_0 \mathbf{X}_L) + \sum_{\ell=1}^L \frac{1}{M_\ell} \mathcal{V}(\tilde{\mathbf{F}}_\ell \mathbf{X}_L - \tilde{\mathbf{F}}_{\ell-1} \mathbf{X}_L). \quad (3.64)$$

Pour une matrice  $\mathbf{M} \in \mathbb{R}^{n_L \times n_L}$ ,

$$\mathcal{V}(\mathbf{M} \mathbf{X}_L) = \mathbb{E}[\|\mathbf{M} \mathbf{X}_L - \mathbb{E}[\mathbf{M} \mathbf{X}_L]\|_{\mathbf{W}_L}^2] = \mathbb{E}[\|\mathbf{M} \dot{\mathbf{X}}_L\|_{\mathbf{W}_L}^2], \quad (3.65)$$

où  $\dot{\mathbf{X}}_L := \mathbf{X}_L - \mathbb{E}[\mathbf{X}_L]$ . On a alors,

$$\mathcal{V}(\mathbf{M} \mathbf{X}_L) = \mathbb{E}[(\mathbf{M} \dot{\mathbf{X}}_L)^T \mathbf{W}_L (\mathbf{M} \dot{\mathbf{X}}_L)] = \text{tr}(\mathbb{E}[\dot{\mathbf{X}}_L^T \mathbf{M}^T \mathbf{W}_L \mathbf{M} \dot{\mathbf{X}}_L]) \quad (3.66)$$

$$= \mathbb{E}[\text{tr}(\dot{\mathbf{X}}_L^T \mathbf{M}^T \mathbf{W}_L \mathbf{M} \dot{\mathbf{X}}_L)] = \mathbb{E}[\text{tr}(\mathbf{M}^T \mathbf{W}_L \mathbf{M} \dot{\mathbf{X}}_L \dot{\mathbf{X}}_L^T)] \quad (3.67)$$

$$= \text{tr}(\mathbf{M}^T \mathbf{W}_L \mathbf{M} \mathbb{E}[\dot{\mathbf{X}}_L \dot{\mathbf{X}}_L^T]) = \text{tr}(\mathbf{M}^T \mathbf{W}_L \mathbf{M} \mathbf{G}) \quad (3.68)$$

$$= \text{tr}((\mathbf{M} \mathbf{G}^{1/2})^T \mathbf{W}_L (\mathbf{M} \mathbf{G}^{1/2})) = \|\mathbf{M} \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2, \quad (3.69)$$

avec  $\mathbf{G} := \mathbb{E}[\dot{\mathbf{X}}_L \dot{\mathbf{X}}_L^T]$  la matrice de covariance de  $\dot{\mathbf{X}}_L$  admettant la décomposition  $\mathbf{G} = \mathbf{G}^{1/2} \mathbf{G}^{1/2}$ ,  $\text{tr}(\cdot)$  l'opérateur trace et  $\|\cdot\|_{F, \mathbf{W}_L} : \mathbf{M} \mapsto \text{tr}(\mathbf{M}^T \mathbf{W}_L \mathbf{M})^{1/2}$  la norme de Frobenius pondérée par la matrice  $\mathbf{W}_L$ . Il est alors possible de réécrire la variance de l'estimateur MLMC de l'équation (3.11)

$$\mathcal{V}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}) = \frac{1}{M_0} \|\tilde{\mathbf{F}}_0 \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2 + \sum_{\ell=1}^L \frac{1}{M_\ell} \|(\tilde{\mathbf{F}}_\ell - \tilde{\mathbf{F}}_{\ell-1}) \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2. \quad (3.70)$$

Pour décomposer cette variance en contributions dans la base de Hartley, on note

$$\tilde{\mathbf{H}}_L = \mathbf{V}_L \mathbf{H}_L \mathbf{V}_L^{-1}, \quad (3.71)$$

et on utilise, pour toute matrice  $\mathbf{M} \in \mathbb{R}^{n_L \times n_L}$ , la relation

$$\|\tilde{\mathbf{H}}_L^T \mathbf{M}\|_{F, \mathbf{W}_L}^2 = \text{tr}(\mathbf{M}^T \tilde{\mathbf{H}}_L \mathbf{W}_L \tilde{\mathbf{H}}_L^T \mathbf{M}) \quad (3.72)$$

$$= \text{tr}(\mathbf{M}^T \tilde{\mathbf{H}}_L \mathbf{V}_L \mathbf{V}_L^T \tilde{\mathbf{H}}_L^T \mathbf{M}) \quad (3.73)$$

$$= \text{tr}(\mathbf{M}^T \mathbf{V}_L \mathbf{H}_L \mathbf{V}_L^{-1} \mathbf{V}_L \mathbf{V}_L^T \mathbf{V}_L^{-T} \mathbf{H}_L^T \mathbf{V}_L^T \mathbf{M}) \quad (3.74)$$

$$= \text{tr}(\mathbf{M}^T \mathbf{V}_L \mathbf{H}_L \mathbf{H}_L^T \mathbf{V}_L^T \mathbf{M}) \quad (3.75)$$

$$= \text{tr}(\mathbf{M}^T \mathbf{V}_L \mathbf{V}_L^T \mathbf{M}) \quad (3.76)$$

$$= \text{tr}(\mathbf{M}^T \mathbf{W}_L \mathbf{M}) \quad (3.77)$$

$$= \|\mathbf{M}\|_{F, \mathbf{W}_L}^2, \quad (3.78)$$

obtenue par orthogonalité des colonnes de  $\mathbf{H}_L$ . On peut alors écrire, en utilisant la relation (3.72) en combinaison avec l'équation (3.70),

$$\mathcal{V}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}) = \frac{1}{M_0} \|\tilde{\mathbf{H}}_L^T \tilde{\mathbf{F}}_0 \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2 + \sum_{\ell=1}^L \frac{1}{M_\ell} \|\tilde{\mathbf{H}}_L^T (\tilde{\mathbf{F}}_\ell - \tilde{\mathbf{F}}_{\ell-1}) \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2. \quad (3.79)$$

Dans le cas de grilles régulières où  $\mathbf{V}_L = n_L^{-1/2} \mathbf{I}_{n_L}$ , on a  $\tilde{\mathbf{H}}_L = \mathbf{H}_L$  et

$$\mathcal{V}(\hat{\boldsymbol{\mu}}_L^{\text{MLMC}}) = \frac{1}{M_0} \|\mathbf{H}_L^T \tilde{\mathbf{F}}_0 \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2 + \sum_{\ell=1}^L \frac{1}{M_\ell} \|\mathbf{H}_L^T (\tilde{\mathbf{F}}_\ell - \tilde{\mathbf{F}}_{\ell-1}) \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2. \quad (3.80)$$

Ces équations mettent en avant que la réduction de la variance de l'estimateur MLMC dépend fortement de la similarité entre les modèles de fidélités successives, c'est-à-dire de la différence  $\tilde{\mathbf{F}}_\ell - \tilde{\mathbf{F}}_{\ell-1}$ . Les modèles  $\tilde{\mathbf{F}}_\ell$  étant construits à partir des modèles originaux  $\mathbf{F}_\ell$  et des opérateurs de transfert, le choix de ces derniers peut avoir un impact important sur la variance de l'estimateur.

### Modèles linéaires, symétriques circulants

On suppose maintenant que les opérateurs linéaires  $\mathbf{F}_\ell$  sont symétriques et circulants pour  $\ell = 0, \dots, L$ . Ces propriétés permettent alors de diagonaliser ces matrices dans la base de Hartley (3.21), comme démontré dans la section A.2. On a, pour  $\ell = 0, \dots, L$ ,

$$\mathbf{F}_\ell = \mathbf{H}_\ell \boldsymbol{\Lambda}_\ell \mathbf{H}_\ell^T, \quad \boldsymbol{\Lambda}_\ell := \text{Diag}(\{\lambda_k^\ell\}_{k=0}^{n_\ell-1}), \quad (3.81)$$

ce qui permet, en utilisant les relations de l'équation (3.58), de réécrire les opérateurs  $\tilde{\mathbf{F}}_\ell$ . Afin de simplifier l'analyse qui suit, on se restreint à deux niveaux de fidélité, et donc à un estimateur MLMC bi-niveaux. On considère les modèles  $\mathbf{F}_0$  et  $\mathbf{F}_1$  définis sur les grilles de taille  $n_0$  et  $n_1 = 2n_0$ . On note  $\mathbf{P} = \mathbf{P}_0^1$  et  $\mathbf{R} = \mathbf{R}_1^0$ , on a  $\tilde{\mathbf{F}}_1 = \mathbf{F}_1 = \mathbf{H}_1 \boldsymbol{\Lambda}_1 \mathbf{H}_1^T$ , et, à l'aide des équations (3.57) et (3.58),

$$\tilde{\mathbf{F}}_0 = \mathbf{P} \mathbf{F}_0 \mathbf{R} = \mathbf{P} \mathbf{H}_0 \boldsymbol{\Lambda}_0 \mathbf{H}_0^T \mathbf{R} = \sqrt{2} \mathbf{H}_1 \mathbf{C}^T \boldsymbol{\Lambda}_0 \mathbf{C} \mathbf{H}_1^T, \quad (3.82)$$

où  $\mathbf{C} = [\text{Diag}(\{c_k\}_{k=0}^{n_0-1}) \text{Diag}(\{-c_{n_0+k}\}_{k=0}^{n_0-1})] \in \mathbb{R}^{n_0 \times n_1}$  est la matrice utilisée dans la section 3.3.2 pour  $\ell = 1$ . On définit la matrice

$$\mathbf{M} = \sqrt{2} \mathbf{C}^T \boldsymbol{\Lambda}_0 \mathbf{C} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix}, \quad (3.83)$$

avec

$$\mathbf{M}_{11} := \sqrt{2} \text{Diag}(\{c_k^2 \lambda_k^0\}_{k=0}^{n_0-1}), \quad (3.84)$$

$$\mathbf{M}_{22} := \sqrt{2} \text{Diag}(\{c_{n_0+k}^2 \lambda_k^0\}_{k=0}^{n_0-1}), \quad (3.85)$$

$$\mathbf{M}_{12} := \sqrt{2} \text{Diag}(\{-c_k c_{n_0+k} \lambda_k^0\}_{k=0}^{n_0-1}) = \mathbf{M}_{21}. \quad (3.86)$$

Donc  $\mathbf{H}_1^\top \tilde{\mathbf{F}}_0 = \mathbf{M} \mathbf{H}_1^\top$  et  $\mathbf{H}_1^\top \mathbf{F}_1 = \Lambda \mathbf{H}_1^\top$ , et l'équation (3.80) devient, pour l'estimateur MLMC à deux niveaux,

$$\mathcal{V}(\hat{\boldsymbol{\mu}}_1^{\text{MLMC}}) = \frac{1}{M_0} \|\mathbf{M} \mathbf{H}_1^\top \mathbf{G}^{1/2}\|_{F, \mathbf{W}_1}^2 + \frac{1}{M_1} \|(\Lambda_1 - \mathbf{M}) \mathbf{H}_1^\top \mathbf{G}^{1/2}\|_{F, \mathbf{W}_1}^2. \quad (3.87)$$

L'hypothèse des modèles linéaires, symétriques et circulants permet ici d'exprimer la différence entre les deux modèles  $\tilde{\mathbf{F}}_1$  et  $\tilde{\mathbf{F}}_0$  en fonction de la différence entre  $\Lambda_1$  et  $\mathbf{M}$ . Il est alors bien plus facile de constater les effets des opérateurs de transfert de grilles sur l'estimateur, les informations relatives aux modèles étant contenues dans leur valeurs propres. Comparés à la matrice composée seulement de la diagonale de  $\mathbf{M}$ , les blocs extra-diagonaux  $\mathbf{M}_{12}$  et  $\mathbf{M}_{21}$  contribuent à l'augmentation de la norme de cette différence, et donc à l'augmentation de la variance de ce terme de correction. Ces blocs extra-diagonaux sont issus de l'opérateur de prolongation qui introduit des composantes petites échelles. Les effets de la restriction se constatent dans le repliement du spectre, les valeurs propres de l'opérateur  $\mathbf{F}_0$  apparaissent deux fois sur la diagonale de  $\mathbf{M}$ . Il est cependant plus compliqué de voir comment réduire la différence entre la diagonale de  $\mathbf{M}$  et  $\Lambda_1$  sans avoir une expression des valeurs propres de  $\mathbf{F}_0$ .

### Opérateurs de Galerkin

On introduit ici l'opérateur de Galerkin qui permet de construire de manière algébrique un modèle sur une grille grossière à partir d'un modèle de plus haute fidélité. Cet opérateur est couramment utilisé dans les méthodes multigrilles [7] et il est ici utilisé pour construire  $\mathbf{F}_0$  à partir de  $\mathbf{F}_1$ ,

$$\mathbf{F}_0 := \frac{1}{2} \mathbf{R} \mathbf{F}_1 \mathbf{P}. \quad (3.88)$$

**Proposition 1.** *L'opérateur  $\mathbf{F}_0$  défini dans l'équation (3.88) est celui qui minimise  $\|\mathbf{F}_1 - \mathbf{P} \mathbf{F}_0 \mathbf{R}\|_{F, \mathbf{W}_1}^2$  pour les opérateurs de transfert fixés dans les équations (3.19) et (3.20) et pour  $\mathbf{W}_1 = n_1^{-1} \mathbf{I}_{n_1}$ .*

**Preuve.** Soient  $n_1, n_0 \in \mathbb{N}$  tels que  $n_1 > n_0 > 0$ , et soient  $\mathbf{A}_1 \in \mathbb{R}^{n_1 \times n_1}$  et  $\mathbf{A}_0 \in \mathbb{R}^{n_0 \times n_0}$  deux opérateurs linéaires. Soient  $\mathbf{R} \in \mathbb{R}^{n_0 \times n_1}$  et  $\mathbf{P} \in \mathbb{R}^{n_1 \times n_0}$  des opérateurs de restriction et de prolongation, respectivement, tels que  $\mathbf{R} \mathbf{R}^\top = \mathbf{I}_{n_0}$  et  $\mathbf{P} = \alpha \mathbf{R}^\top$ , pour un  $\alpha \in \mathbb{R}$  fixé. On cherche à obtenir l'opérateur  $\mathbf{A}_0^*$  qui minimise  $\|\mathbf{A}_1 - \mathbf{P} \mathbf{A}_0 \mathbf{R}\|_{F, \mathbf{W}_1}^2$  parmi tous les opérateurs linéaires  $\mathbf{A}_0 \in \mathbb{R}^{n_0 \times n_0}$ . On se restreint au cas  $\mathbf{W}_1 = w \mathbf{I}_{n_1}$ , avec  $w > 0$ , pour avoir

$$\mathbf{A}_0^* := \arg \min_{\mathbf{A}_0 \in \mathbb{R}^{n_0 \times n_0}} \|\mathbf{A}_1 - \mathbf{P} \mathbf{A}_0 \mathbf{R}\|_{F, \mathbf{W}_1}^2 = \arg \min_{\mathbf{A}_0 \in \mathbb{R}^{n_0 \times n_0}} \|\mathbf{A}_1 - \mathbf{P} \mathbf{A}_0 \mathbf{R}\|_F^2, \quad (3.89)$$

où  $\|\cdot\|_F$  correspond à la norme de Frobenius. On commence par écrire

$$\|\mathbf{A}_1 - \mathbf{P}\mathbf{A}_0\mathbf{R}\|_F^2 = \|\mathbf{A}_1 - \mathbf{P}\mathbf{A}_0\mathbf{R}\|_F^2 \quad (3.90)$$

$$= \|\text{vec}(\mathbf{A}_1 - \mathbf{P}\mathbf{A}_0\mathbf{R})\|_2^2 \quad (3.91)$$

$$= \|\text{vec}(\mathbf{A}_1) - \text{vec}(\mathbf{P}\mathbf{A}_0\mathbf{R})\|_2^2 \quad (3.92)$$

$$= \|\text{vec}(\mathbf{A}_1) - (\mathbf{R}^T \otimes \mathbf{P}) \text{vec}(\mathbf{A}_0)\|_2^2, \quad (3.93)$$

où  $\text{vec}(\cdot)$  est l'opérateur de vectorisation d'une matrice qui empile ses colonnes les unes sur les autres. L'équation (3.89) est ainsi écrit sous la forme d'un problème aux moindres carrés ordinaires dont la solution est

$$\text{vec}(\mathbf{A}_0^*) = \left( (\mathbf{R}^T \otimes \mathbf{P})^T (\mathbf{R}^T \otimes \mathbf{P}) \right)^{-1} (\mathbf{R}^T \otimes \mathbf{P})^T \text{vec}(\mathbf{A}_1) \quad (3.94)$$

$$= \left( (\mathbf{R} \otimes \mathbf{P}^T) (\mathbf{R}^T \otimes \mathbf{P}) \right)^{-1} (\mathbf{R} \otimes \mathbf{P}^T) \text{vec}(\mathbf{A}_1) \quad (3.95)$$

$$= (\mathbf{R}\mathbf{R}^T \otimes \mathbf{P}^T\mathbf{P})^{-1} \text{vec}(\mathbf{P}^T\mathbf{A}_1\mathbf{R}^T) \quad (3.96)$$

$$= (\mathbf{I}_{n_0} \otimes \alpha^2 \mathbf{I}_{n_0})^{-1} \text{vec}((\alpha\mathbf{R})\mathbf{A}_1(\alpha^{-1}\mathbf{P})) \quad (3.97)$$

$$= \alpha^{-2} (\mathbf{I}_{n_0} \otimes \mathbf{I}_{n_0}) \text{vec}(\mathbf{R}\mathbf{A}_1\mathbf{P}) \quad (3.98)$$

$$= \alpha^{-2} \text{vec}(\mathbf{R}\mathbf{A}_1\mathbf{P}), \quad (3.99)$$

où  $\otimes$  désigne le produit de Kronecker. Ce résultat donne bien  $\mathbf{A}_0^* = \alpha^{-2}\mathbf{R}\mathbf{A}_1\mathbf{P}$ . Les opérateurs de transfert des équations (3.19) et (3.20) suivent la relation  $\mathbf{P} = \sqrt{2}\mathbf{R}^T$ , on a  $\alpha = \sqrt{2}$  et  $\mathbf{A}_0^* = (1/2)\mathbf{R}\mathbf{A}_1\mathbf{P}$ .

Quand la matrice de covariance de la variable aléatoire centrée  $\dot{\mathbf{X}}_L$  est la matrice identité, (3.88) minimise la variance du terme de correction de l'équation (3.70). Il est à noter que l'opérateur de Galerkin est rarement utilisé en pratique car il doit être calculé et stocké explicitement. Il permet, cependant, dans des développement théoriques comme celui ci, de s'assurer d'avoir la définition optimale d'un opérateur basse fidélité. En utilisant les identités (3.57) et (3.58), on a

$$\mathbf{F}_0 = \frac{1}{2}\mathbf{R}\mathbf{F}_1\mathbf{P} = \frac{1}{2}\mathbf{R}\mathbf{H}_1\mathbf{\Lambda}_1\mathbf{H}_1^T\mathbf{P} = \frac{\sqrt{2}}{2}\mathbf{H}_0\mathbf{C}\mathbf{\Lambda}_1\mathbf{C}^T\mathbf{H}_0^T. \quad (3.100)$$

On simplifie à nouveau les expressions suivantes en omettant les indices de début et de fin des suites qui vont, ici, toujours de  $k = 0$  à  $k = n_0 - 1$ . On a

$$\mathbf{C}\mathbf{\Lambda}_1\mathbf{C}^T = \begin{bmatrix} \text{Diag}(\{c_k\}) \\ \text{Diag}(\{-c_{n_0+k}\}) \end{bmatrix}^T \begin{bmatrix} \text{Diag}(\{\lambda_k^1\}) & \mathbf{0} \\ \mathbf{0} & \text{Diag}(\{\lambda_{n_0+k}^1\}) \end{bmatrix} \begin{bmatrix} \text{Diag}(\{c_k\}) \\ \text{Diag}(\{-c_{n_0+k}\}) \end{bmatrix} \quad (3.101)$$

$$= \text{Diag}(\{c_k^2\lambda_k^1\}) + \text{Diag}(\{c_{n_0+k}^2\lambda_{n_0+k}^1\}). \quad (3.102)$$

La matrice  $\mathbf{\Lambda}_0 := \sqrt{2}^{-1}\mathbf{C}\mathbf{\Lambda}_1\mathbf{C}^T$  est donc la matrice diagonale des valeurs propres de  $\mathbf{F}_0$  et on peut exprimer les valeurs propres  $\{\lambda_k^0\}_{k=0}^{n_0-1}$  de  $\mathbf{F}_0$  en fonction de celles de  $\mathbf{F}_1$ ,

$$\lambda_k^0 = \frac{1}{\sqrt{2}}(c_k^2\lambda_k^1 + c_{n_0+k}^2\lambda_{n_0+k}^1), \quad \forall k = 0, \dots, n_0 - 1. \quad (3.103)$$

Les blocs de la matrice  $\mathbf{M}$ , définie dans l'équation (3.83), deviennent alors

$$\mathbf{M}_{11} := \text{Diag}(\{c_k^4 \lambda_k^1 + c_k^2 c_{n_0+k}^2 \lambda_{n_0+k}^1\}_{k=0}^{n_0-1}), \quad (3.104)$$

$$\mathbf{M}_{22} := \text{Diag}(\{c_{n_0+k}^4 \lambda_k^1 + c_k^2 c_{n_0+k}^2 \lambda_k^1\}_{k=0}^{n_0-1}) = \text{Diag}(\{c_k^4 \lambda_k^1 + c_{k-n_0}^2 c_k^2 \lambda_{k-n_0}^1\}_{k=n_0}^{2n_0-1}), \quad (3.105)$$

$$\mathbf{M}_{12} := \text{Diag}(\{-c_k^3 c_{n_0+k} \lambda_k^1 - c_k c_{n_0+k}^3 \lambda_{n_0+k}^1\}_{k=0}^{n_0-1}) = \mathbf{M}_{21}. \quad (3.106)$$

La réécriture du bloc  $\mathbf{M}_{22}$  en utilisant la relation  $-c_{n_0+k} = c_{n_0+k-2n_0} = c_{k-n_0}$  permet de mieux comparer la diagonale de  $\mathbf{M}$  à  $\Lambda_1$ . On peut, en effet, constater que l'écart entre la diagonale de  $\mathbf{M}$  et  $\Lambda_1$  se situe, premièrement, dans le facteur réducteur  $c_k^4 \leq 1$  associé à la  $k$ -ième valeur propre, et deuxièmement, dans le terme additif composé d'une valeur propre complémentaire, heureusement réduit par un facteur  $c_k^2 c_{n_0+k}^2 < 1$  pour tout  $k = 0, \dots, 2n_0 - 1$ . Les blocs extra-diagonaux, quant à eux, contribuent à augmenter la norme de la différence et l'on utilise donc directement leur expression donnée dans l'équation (3.85). Afin de mieux visualiser l'écart entre  $\mathbf{M}$  et  $\Lambda_1$ , la figure 3.7 représente les différents coefficients sous forme graphique.

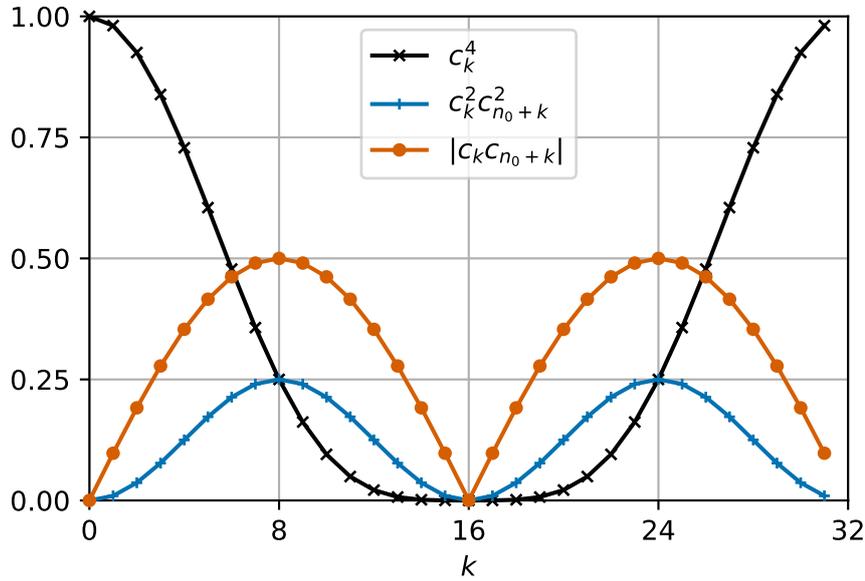


FIGURE 3.7 – Les facteurs devant les valeurs propres dans la matrice  $\mathbf{M}$  en fonction de  $k = 0, \dots, n_1 - 1$ . La courbe noire correspond aux facteurs devant la bonne valeur propre sur la diagonale de  $\mathbf{M}$ , tandis que la courbe bleue est associée aux facteurs devant le terme parasite des valeurs propres complémentaires. Finalement la courbe orange correspond aux facteurs des blocs extra-diagonaux dans leur écriture en fonction de  $\lambda^0$ .

Pour que  $\mathbf{M}$  soit égale à  $\Lambda_1$  il faudrait que la courbe noire soit constante et égale à 1, et que les courbes bleue et orange soient nulles. On remarque donc que, pour  $k$  proche de 0 ou de  $n_1 - 1$ , on s'approche d'un tel résultat car les coefficients  $c_k^4$  devant la  $k$ -ième valeur propre sont proches de 1, tandis que les coefficients des valeurs propres complémentaires  $n_0 + k$  et des termes extra-diagonaux sont proches de 0. Pour  $k = 0$  la représentation est même parfaite car  $c_0^4 = 1$  tandis que  $c_k^2 c_{n_0+k}^2 = c_k c_{n_0+k} = 0$ . Les composantes associées aux grandes échelles (basses fréquences) qui correspondent aux  $k$  proches de 0 ou de  $n_1 - 1$  sont donc plutôt bien représentées. Cependant la rapide décroissance des facteurs  $c_k^4$  quand  $k$  tend vers  $n_0$  conduit à une représentation

fortement atténuée des moyennes et petites échelles. L'amplitude des composantes parasites devient alors bien plus importante relativement aux composantes correctes, aboutissant à des représentations très éloignées des valeurs propres de  $\Lambda_1$ . Excepté pour la valeur propre  $\lambda_0^1$  qui est parfaitement représentée, toutes les autres composantes sont erronées avec une erreur qui croît quand  $k$  tend vers  $n_0$ . Les petites échelles, qui ne peuvent pas être représentées sur la grille grossière, sont alors très mal estimées.

### 3.4 Illustration 1D

Après avoir exprimé les effets des opérateurs de prolongation et de restriction choisis ici sur la base de Hartley, on a pu exprimer leur impact sur la variance d'un estimateur MLMC à deux niveaux dans le cas de modèles linéaires, symétriques et circulants. Ces opérateurs de transfert introduisent des composantes parasites et créent un repliement du spectre qui apporte des erreurs dans la représentation de toutes les échelles (sauf la composante constante associée à la première valeur propre), et, plus spécifiquement, aux petites échelles. Dans cette section, on illustre ces différents phénomènes et leur impact sur la variance de l'estimateur grâce à des expériences numériques sur un modèle jouet.

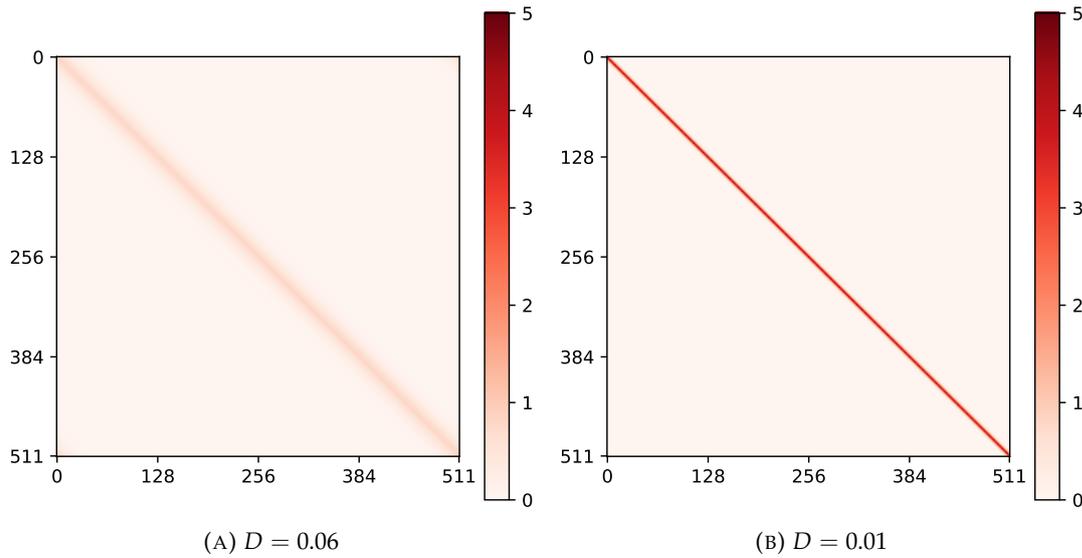
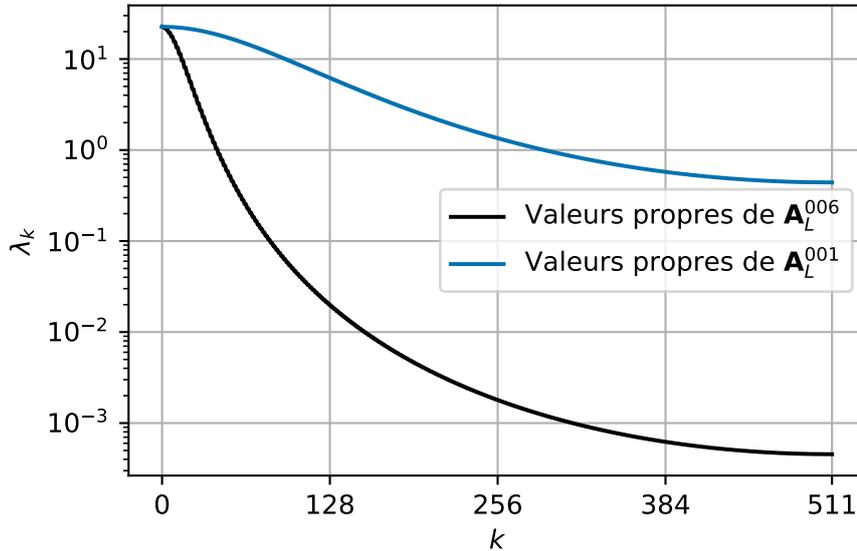
#### 3.4.1 Présentation du problème

Pour illustrer les résultats précédents on se place dans le même contexte. L'expérience numérique se base donc sur une discrétisation, par  $L + 1$  grilles centrées sur les cellules, du domaine  $\mathcal{D} = [0, 1]$ . Les grilles sont uniformes et composées de  $n_\ell$  cellules, chacune de taille  $n_\ell^{-1}$ , pour  $\ell = 0, \dots, L$ . Leur matrice de Gram associée est  $\mathbf{W}_\ell = n_\ell^{-1} \mathbf{I}_{n_\ell}$ . On considère différentes hiérarchies pour  $L \in \{0, \dots, 5\}$  en fixant la taille de la grille la plus fine à  $n_L = 512$ . La taille des grilles plus grossières est obtenue avec un facteur de réduction fixé  $n_{\ell+1}/n_\ell = 2$  pour  $\ell = 0, \dots, L - 1$ , ce qui donne  $n_\ell = 2^{\ell-L} n_L$ . Les opérateurs de prolongation et de restriction entre deux grilles successives sont ceux définis dans les équations (3.19) et (3.20), et les opérateurs de transfert entre la grille fine et une autre grille arbitraire sont définis dans les équations (3.4) et (3.5). La matrice de Hartley  $\mathbf{H}_L$  utilisée pour décomposer la variance des estimateurs est définie dans l'équation (3.21).

Le simulateur  $f_L$  sur la grille fine est choisi linéaire, symétrique et circulant et est donc représenté par une matrice  $\mathbf{A}_L \in \mathbb{R}^{512 \times 512}$ . La construction de cette matrice sera détaillée dans le chapitre 5 à partir, notamment, de l'équation (5.18). On se contentera ici de mentionner que cette matrice correspond à un facteur de la décomposition d'une matrice de covariance  $\mathbf{L}_L = \mathbf{A}_L \mathbf{A}_L^T$  avec des conditions aux bords périodiques, et donc, pour  $\mathbf{X}_L \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{I}_{n_L})$ ,  $\mathbf{A}_L \mathbf{X}_L \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{L}_L)$ . Un paramètre  $D \in \mathbb{R}$  est utilisé pour représenter la longueur de corrélation de cet opérateur. Pour cette illustration, deux longueurs de corrélation sont considérées,  $D = 0.06 \approx 30n_L^{-1}$  et  $D = 0.01 \approx 5n_L^{-1}$ . On notera  $\mathbf{A}_L^{006}$  (resp.  $\mathbf{A}_L^{001}$ ) la matrice construite à partir de la longueur de corrélation  $D = 0.06$  (resp.  $D = 0.01$ ). La figure 3.8 présente les deux matrices  $\mathbf{A}_L^{006}$  et  $\mathbf{A}_L^{001}$ .

Ces opérateurs étant, par construction, symétriques et circulants, ils sont diagonalisables dans la base de Hartley  $\mathbf{H}_L$ . La figure 3.9 représente les valeurs propres des matrices  $\mathbf{A}_L^{006}$  et  $\mathbf{A}_L^{001}$  dans la base de Hartley ordonnée suivant les colonnes de  $\mathbf{H}_L^*$  (cf. équation (3.23)).

Une longueur de corrélation plus élevée réduit les valeurs propres associées aux petites échelles, aboutissant à un champ contenant principalement des composantes


 FIGURE 3.8 – Représentation des matrices  $\mathbf{A}_L^{006}$  et  $\mathbf{A}_L^{001}$ .

 FIGURE 3.9 – Valeurs propres  $\lambda_k$  des matrices  $\mathbf{A}_L^{006}$  (noir) et  $\mathbf{A}_L^{001}$  (bleu) dans la base de Hartley réordonnée  $\mathbf{H}_L^*$ .

grandes échelles. À l'inverse, une longueur de corrélation plus petite augmente les valeurs propres associées aux moyennes et petites échelles qui deviennent alors bien plus significatives.

Le problème étudié est celui de l'estimation de  $\mathbb{E}[\mathbf{A}_L \mathbf{X}_L]$  où  $\mathbf{X}_L \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{I}_{n_L})$  est un vecteur aléatoire de taille  $n_L$  suivant une loi normale multivariée. Donc, dans notre cas, la matrice  $\mathbf{G}$  définie dans la section 3.3.3 est telle que  $\mathbf{G} = \mathbb{E}[\mathbf{X}_L \mathbf{X}_L^T] = \mathbf{I}_{n_L}$ . Par linéarité de l'espérance, on a immédiatement  $\mathbb{E}[\mathbf{A}_L \mathbf{X}_L] = \mathbf{0}_{n_L}$ . On connaît donc l'espérance exacte que l'on souhaite estimer, ce qui permet de facilement calculer la MSE des estimateurs MC et MLMC. L'estimateur MLMC se base sur les simulateurs  $f_\ell$  définis sur les grilles grossières comme les opérateurs de Galerkin, i.e.,

$$f_\ell = 2^{\ell-L} \mathbf{R}_L^\ell \mathbf{A}_L \mathbf{P}_\ell^L, \quad \ell = 0, \dots, L-1, \quad (3.107)$$

et les opérateurs utilisés dans l'estimateur (3.7) sont  $\tilde{f}_\ell = 2^{\ell-L} \mathbf{P}_\ell^L \mathbf{R}_L^\ell \mathbf{A}_L \mathbf{P}_\ell^L \mathbf{R}_L^\ell$ . Pour

comparer les différents estimateurs, un budget de calcul  $\eta$  est fixé pour toutes les estimations. On note  $\mathcal{C}_\ell$  le coût de calcul moyen nécessaire à une évaluation de  $\tilde{f}_\ell$  en  $\mathbf{X}_L$  et on suppose qu'un tel coût est linéaire avec la taille de la grille du niveau  $\ell$ , i.e.,  $\mathcal{C}_\ell = \mathcal{O}(n_\ell)$ . Cette hypothèse se justifie ici par les simulateurs  $\tilde{f}_\ell$  qui correspondent à l'application d'une matrice creuse dont le nombre d'entrées non-nulles est proportionnel à la taille  $n_\ell$  de la matrice  $\mathbf{A}_\ell$ . Le facteur de réduction  $n_{\ell+1}/n_\ell$  fixé à 2 donne donc le même facteur de coût  $\mathcal{C}_{\ell+1}/\mathcal{C}_\ell = 2$  pour  $\ell = 0, \dots, L$ . On fixe le budget total  $\eta = 100\mathcal{C}_L$ . L'estimateur MC requiert donc 100 applications de  $\mathbf{A}_L$ , tandis que pour le MLMC on cherche une allocation optimale du nombre d'applications sur chaque niveau  $\{M_\ell\}_{\ell=0}^L$  pour réduire la variance de l'estimateur sous la contrainte du budget total imposé. Cette allocation optimale est donnée, selon [52], par

$$M_\ell = \left\lfloor \frac{\eta}{\mathcal{S}_L} \sqrt{\frac{\mathcal{V}_\ell}{\mathcal{C}_\ell + \mathcal{C}_{\ell-1}}} \right\rfloor^+, \quad \mathcal{S}_L := \sum_{\ell=0}^L \sqrt{\mathcal{V}_\ell(\mathcal{C}_\ell + \mathcal{C}_{\ell-1})}, \quad (3.108)$$

avec  $\mathcal{V}_\ell := \mathcal{V}(\mathbf{Y}_\ell - \mathbf{Y}_{\ell-1})$  et  $\lfloor \cdot \rfloor^+ := \max(1, \lfloor \cdot \rfloor)$  où  $\lfloor \cdot \rfloor$  est la fonction partie entière. L'équation (3.108) est valable pour  $\ell = 0, \dots, L$  avec, par convention,  $\mathcal{C}_{-1} = 0$ . Il faut noter que cette allocation optimale peut aboutir à un budget de l'estimateur MLMC qui dépasse le budget total imposé  $\eta$  d'un coût additionnel inférieur à  $\mathcal{C}_L$ . Ce coût additionnel est négligeable si  $\eta \gg \mathcal{C}_L$  comme c'est le cas ici. La valeur des  $\mathcal{V}_\ell$  n'est pas connue mais est estimée avec un échantillon pilote de taille 1000.

### 3.4.2 Résultats

On commence par étudier le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  avec  $\mathbf{X}_L \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{I}_{n_L})$ . La figure 3.10 présente les quantités

$$\mathcal{E}_\ell = \|\mathbb{E}[\mathbf{Y}_\ell - \mathbf{Y}_{\ell-1}]\|_{\mathbf{W}_L} \quad \text{et} \quad \mathcal{V}_\ell = \mathbb{E}[\|\mathbf{Y}_\ell - \mathbf{Y}_{\ell-1}\|_{\mathbf{W}_L}^2] - \|\mathbb{E}[\mathbf{Y}_\ell - \mathbf{Y}_{\ell-1}]\|_{\mathbf{W}_L}^2 \quad (3.109)$$

pour  $\ell = 1, \dots, L$ , avec  $\mathcal{E}_0 = \|\mathbb{E}[\mathbf{Y}_0]\|_{\mathbf{W}_L}$  et  $\mathcal{V}_0 = \mathbb{E}[\|\mathbf{Y}_0\|_{\mathbf{W}_L}^2] - \|\mathbb{E}[\mathbf{Y}_0]\|_{\mathbf{W}_L}^2$ . Elle présente également une régression de ces quantités pour obtenir une estimation des paramètres  $\alpha$  et  $\beta$  du théorème 1.

Pour ce problème, on a clairement les relations

$$\mathcal{E}_\ell \lesssim n_\ell^{-\alpha}, \quad \mathcal{V}_\ell \lesssim n_\ell^{-\beta}, \quad (3.110)$$

supposées dans le théorème de convergence de l'estimateur MLMC avec, ici,  $\alpha \approx 0.881$  et  $\beta \approx 1.891$ . On rappelle que l'on a supposé un modèle de coût linéaire pour l'application des différents simulateurs, on a donc  $\mathcal{C}_\ell \lesssim n_\ell^\gamma$  avec  $\gamma = 1$ . On se retrouve bien dans le cas favorable  $\beta > \gamma$ .

Avec les quantités  $\mathcal{V}_\ell$  estimées, l'allocation optimale peut être calculée et l'estimateur MLMC appliqué. La figure 3.11a présente l'allocation optimale de trois estimateurs MLMC, un MLMC à 2 niveaux ( $L = 1$ ), 4 niveaux ( $L = 3$ ) et 6 niveaux ( $L = 5$ ). Les variances totales  $\mathcal{V}(\hat{\mu}_L^{\text{MLMC}})$  de ces estimateurs sont tracées sur la figure 3.11b et comparées à la variance de l'estimateur MC utilisant le même budget  $\eta$ .

Les estimateurs MLMC ont bien une variance plus faible que l'estimateur MC pour un même budget de calcul. L'utilisation de deux niveaux de fidélité permet de réduire la variance de plus de 40% et jusqu'à plus de 70% pour six niveaux. L'allocation optimale montre ici que la variance  $\mathcal{V}_L$  entre les deux niveaux les plus fins  $L$  et  $L - 1$  est très faible et nécessite donc très peu de membres. On rappelle cependant que la

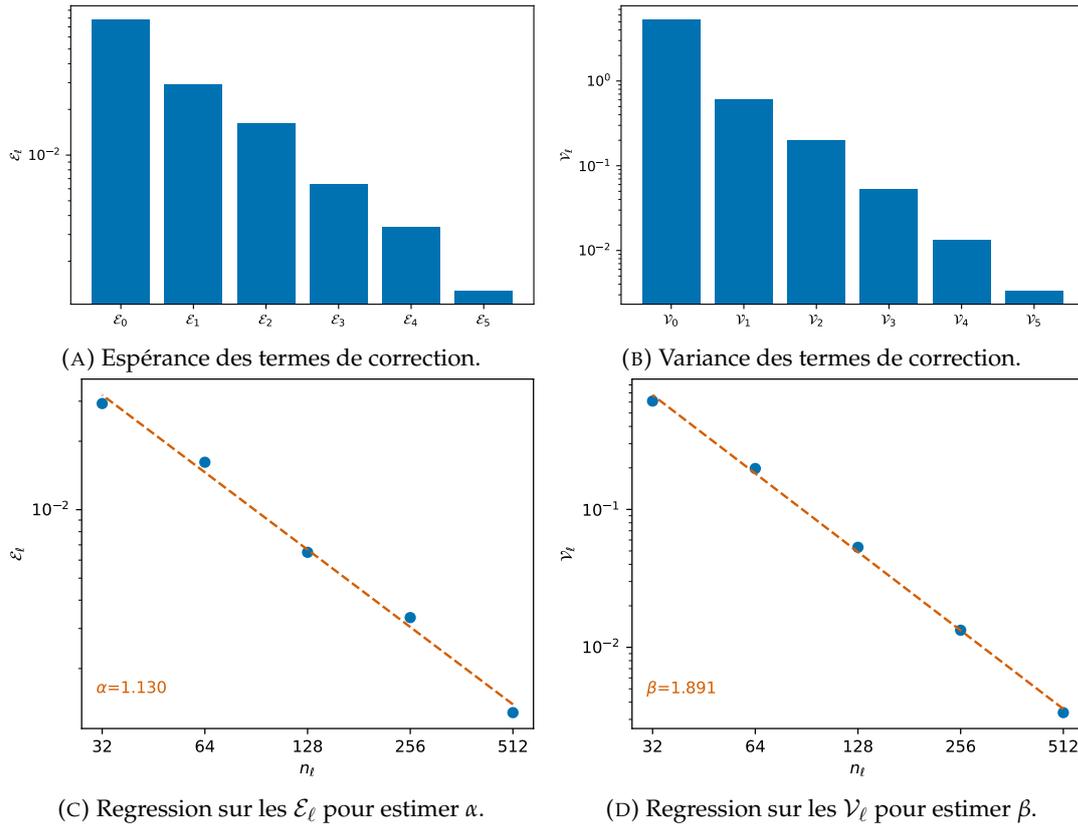


FIGURE 3.10 – Estimation des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  pour  $\ell = 0, \dots, L$  (haut) et regression sur ces quantites afin d'estimer  $\alpha$  et  $\beta$  (bas) pour le probleme d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Les estimations sont faites à partir d'un échantillon de taille 1000.

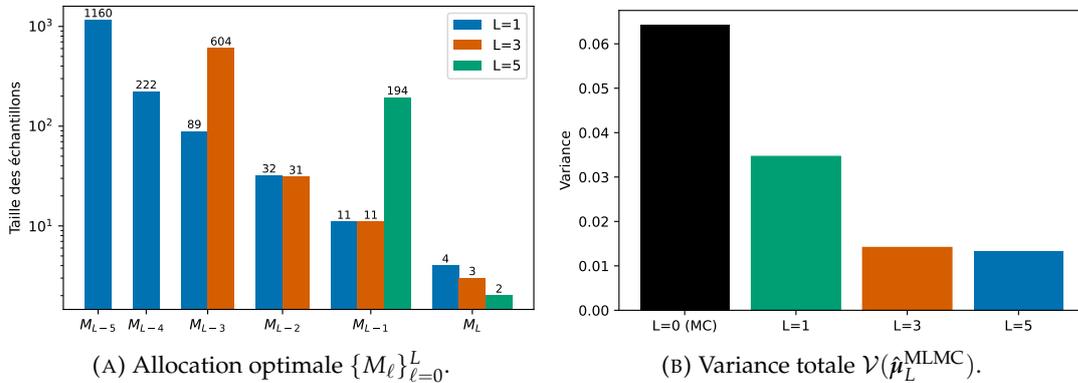


FIGURE 3.11 – Allocation optimale des échantillons sur les différents niveaux et la variance totale résultante des estimateurs MC ( $L = 0$ ) et de différents estimateurs MLMC ( $L \in \{1, 3, 5\}$ ) pour le probleme d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

variance totale définie dans l'équation (3.11) présentée ici est une quantité intégrée et qu'elle ne donne aucune indication sur les échelles spatiales. On sait que la variance totale est réduite mais on ne peut pas savoir si la variance de toutes les composantes est réduite ou seulement celle de quelques composantes précises. La décomposition dans la base de Hartley de la MSE (ici, de la variance) donnée dans l'équation (3.27)

donne accès à la variance de chaque composante d'échelle avec le vecteur  $\boldsymbol{\nu}$  de la variance spectrale (cf. équations (3.35) et (3.39)). On définit également le vecteur de variance cumulée  $\boldsymbol{\nu}^{\text{cml}} = (\nu_k^{\text{cml}})_{k=0}^{n_L-1}$  tel que  $\nu_k^{\text{cml}} = \sum_{k'=0}^k \nu_{k'}$ . La variance totale est alors donnée par  $\nu_{n_L-1}^{\text{cml}} = \|\boldsymbol{\nu}\|_1$ . Dans les figures qui suivent, la décomposition est faite avec la matrice de Hartley réordonnée  $\mathbf{H}_L^*$  pour que les variances soient classées des plus grandes échelles aux plus petites. Les vecteurs de variance spectrale  $\boldsymbol{\nu}$  et de variance cumulée  $\boldsymbol{\nu}^{\text{cml}}$  pour l'estimateur MC et les trois estimateurs MLMC utilisés plus haut sont affichés sur les figures 3.12a et 3.12b.

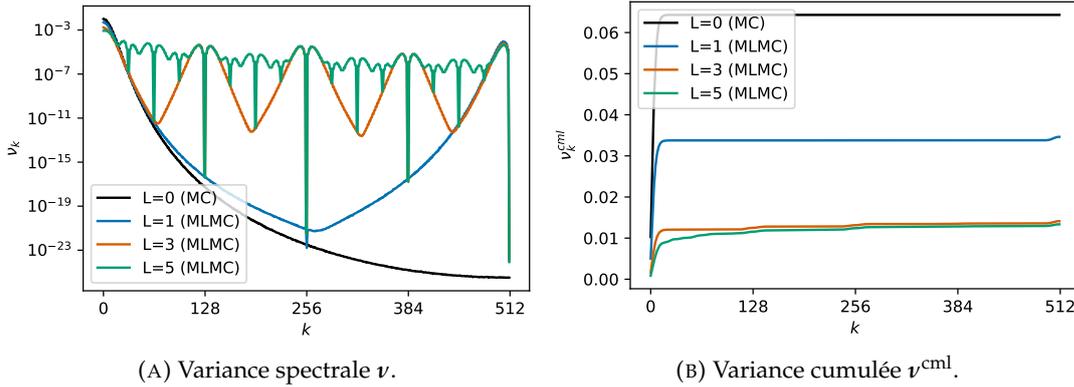


FIGURE 3.12 – Variance spectrale et cumulée de l'estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC ( $L \in \{1, 3, 5\}$ ) pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

On constate, premièrement, que pour l'estimateur MC, la variance associée aux grandes échelles est bien plus importante que celle des petites échelles. La contribution des petites échelles à la variance totale est négligeable comme on peut le constater sur le tracé de la variance cumulée. On remarque également que sa variance spectrale est très similaire au profil des valeurs propres de l'opérateur  $\mathbf{A}_L^{006}$  de la figure 3.9. En effet, pour un estimateur MC basé sur un échantillon de taille  $M$ , on peut exprimer sa variance spectrale (cf. équation (3.35)) comme

$$\nu_k = \frac{1}{n_L} \mathbb{V}[(\mathbf{h}_k^L)^T \hat{\boldsymbol{\mu}}_L] = \frac{1}{Mn_L} \mathbb{V}[(\mathbf{h}_k^L)^T \mathbf{F}_L \mathbf{X}_L] \quad (3.111)$$

$$= \frac{1}{Mn_L} \mathbb{E}[(\mathbf{h}_k^L)^T \mathbf{F}_L \dot{\mathbf{X}}_L]^2 = \frac{1}{Mn_L} \mathbb{E}[(\mathbf{h}_k^L)^T \mathbf{F}_L \dot{\mathbf{X}}_L \dot{\mathbf{X}}_L^T \mathbf{F}_L^T \mathbf{h}_k^L] \quad (3.112)$$

$$= \frac{1}{Mn_L} (\mathbf{h}_k^L)^T \mathbf{F}_L \mathbb{E}[\dot{\mathbf{X}}_L \dot{\mathbf{X}}_L^T] \mathbf{F}_L^T \mathbf{h}_k^L = \frac{1}{Mn_L} (\mathbf{h}_k^L)^T \mathbf{F}_L^2 \mathbf{h}_k^L \quad (3.113)$$

$$= \frac{1}{Mn_L} (\mathbf{h}_k^L)^T \mathbf{H}_L \boldsymbol{\Lambda}_L \mathbf{H}_L^T \mathbf{H}_L \boldsymbol{\Lambda}_L \mathbf{H}_L^T \mathbf{h}_k^L \quad (3.114)$$

$$= \frac{1}{Mn_L} (\mathbf{h}_k^L)^T \mathbf{H}_L \boldsymbol{\Lambda}_L^2 ((\mathbf{h}_k^L)^T \mathbf{H}_L)^T \quad (3.115)$$

$$= \frac{1}{Mn_L} \mathbf{e}_k^T \boldsymbol{\Lambda}_L^2 \mathbf{e}_k \quad (3.116)$$

$$= \frac{1}{Mn_L} (\lambda_k^L)^2. \quad (3.117)$$

Deuxièmement, pour l'estimateur MLMC à deux niveaux, la variance spectrale pour la première moitié du spectre apparaît similaire à celle du MC, elle se différencie

cependant sur la seconde moitié du spectre où l'on observe une variance bien plus importante pour les petites échelles. L'échelle logarithmique de la variance spectrale peut donner l'impression que la variance totale du MLMC est plus élevée que celle du MC mais la variance cumulée permet de voir que la contribution des petites échelles à la variance totale reste négligeable comparée à la contribution des grandes échelles. La réduction de variance du MLMC par rapport au MC provient alors essentiellement d'une meilleure estimation des grandes échelles. Pour les estimateurs MLMC avec un plus grand nombre de niveaux, la détérioration de la variance spectrale comparée au MC s'étend sur quasiment tout le spectre avec seulement les quelques premières composantes qui ont une variance plus faible. Cette réduction de la variance sur les grandes échelles est plus importante pour  $L \in \{3, 5\}$  comparé à  $L = 1$ . La variance associée aux grandes échelles reste, cependant, largement dominante dans la variance totale qui est donc réduite par rapport à celle du MC. La contribution des autres échelles à la variance devient de plus en plus visible dans la variance cumulée avec l'ajout de niveaux, et donc de moins en moins négligeable.

L'analyse spectrale réalisée dans la section 3.3 a permis de conclure que les opérateurs de transfert, nécessaires pour le passage entre les grilles, conduisent à une mauvaise estimation des composantes petites échelles, non représentables sur les grilles grossières, à cause de l'apparition de fréquences parasites. Cette mauvaise représentation se traduit, dans la variance de l'estimateur, par la différence entre les valeurs propres de l'opérateur  $\mathbf{A}_L^{006}$  et la matrice  $\mathbf{M}$  définie dans la section 3.3.3. La norme de cette différence, et donc son impact sur la variance totale, dépend des valeurs propres du modèle fin. Dans cette illustration avec l'opérateur  $\mathbf{A}_L^{006}$ , les valeurs propres associées aux petites échelles sont suffisamment faibles pour que leur contribution à la variance totale soit peu élevée, voire négligeable. On répète alors la même expérience avec l'opérateur  $\mathbf{A}_L^{001}$  dont les valeurs propres, affichées sur la figure 3.9, sont plus élevées, notamment celles associées aux fines échelles. La figure 3.13 présente les quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  ainsi que l'estimation des quantités  $\alpha$  et  $\beta$  du théorème 1.

Les relations de décroissance des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$ , supposées dans le théorème de convergence, ne semblent pas vérifiées ici pour  $\ell = 0, \dots, L$ . Les deux niveaux les plus grossiers ( $\ell = L - 5$  et  $\ell = L - 4$ ), spécifiquement, ont des valeurs  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  plus basses que celles attendues, ce qui indique que ces niveaux risquent de détériorer la réduction de variance du MLMC. Cela peut s'expliquer par le fait que, sur les grilles de taille 16 et 32, une majorité du spectre de  $\mathbf{A}_L^{001}$  n'est pas représentable. Les valeurs propres non-représentables ayant une amplitude bien plus élevées pour  $\mathbf{A}_L^{001}$  que pour  $\mathbf{A}_L^{006}$  (cf. figure 3.9), cela peut correspondre à un pourcentage plus élevé de la trace de la matrice. Les opérateurs de Galerkin associés à ces niveaux contiennent également des applications successives des opérateurs de transfert dont les effets négatifs sont alors amplifiés. Les termes parasites introduits par les opérateurs de transfert, et apparaissant dans la matrice  $\mathbf{M}$  des équations (3.104) à (3.106), dépendent des valeurs propres de  $\mathbf{A}_L^{001}$  et sont, ici, plus importants. L'allocation optimale des différents estimateurs MLMC ainsi que leur variance est présentée dans la figure 3.14.

L'allocation optimale des différents estimateurs conduit, ici, à des échantillons plus grands sur les niveaux les plus fins où une plus grande partie du spectre de  $\mathbf{A}_L^{001}$  est représentable. On constate cependant que, malgré cette allocation, le MLMC à 6 niveaux a une variance supérieure à celle du simple MC. La non-vérification des hypothèses du théorème de convergence ne permet pas de garantir la réduction de variance du MLMC comparé au MC. Les estimateurs MLMC à 2 et 4 niveaux ont une variance plus faible que celle du MC mais la réduction est bien plus faible que dans l'expérience précédente. L'ajout des grilles de taille 128 ou moins détériore

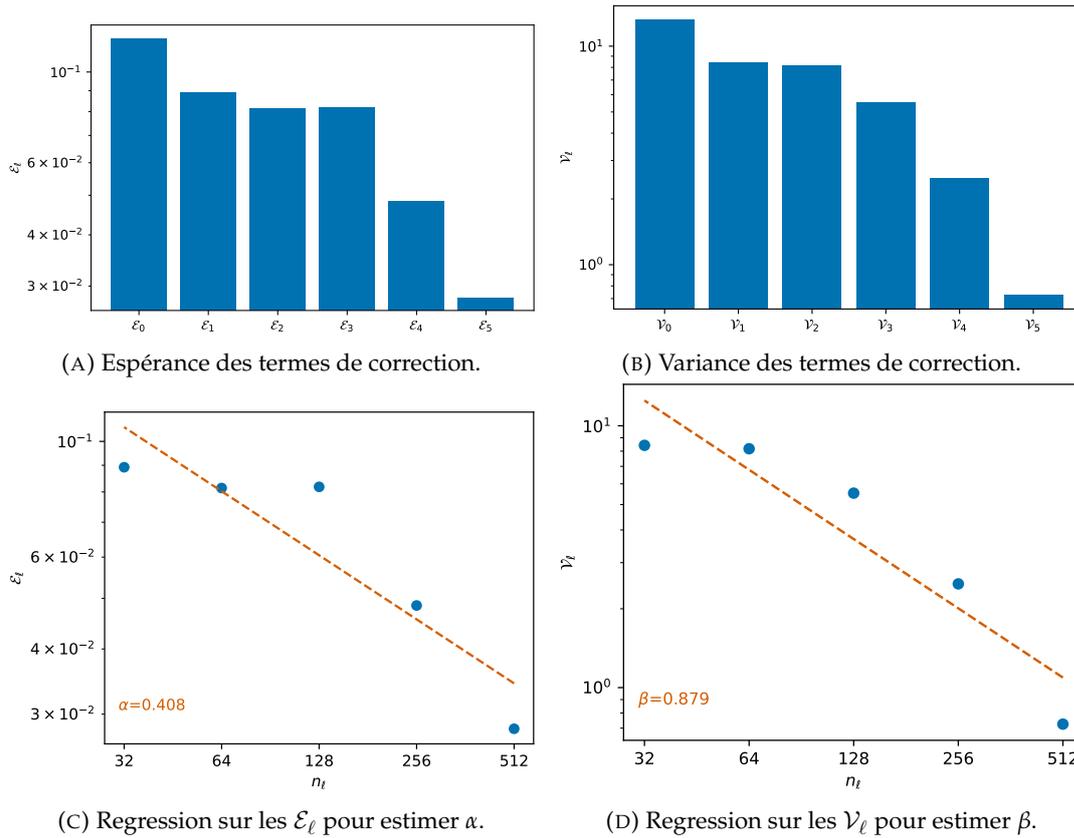


FIGURE 3.13 – Estimation des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  pour  $\ell = 0, \dots, L$  (haut) et regression sur ces quantités afin d’estimer  $\alpha$  et  $\beta$  (bas) pour le problème d’estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ . Les estimations sont faites à partir d’un échantillon de taille 1000.

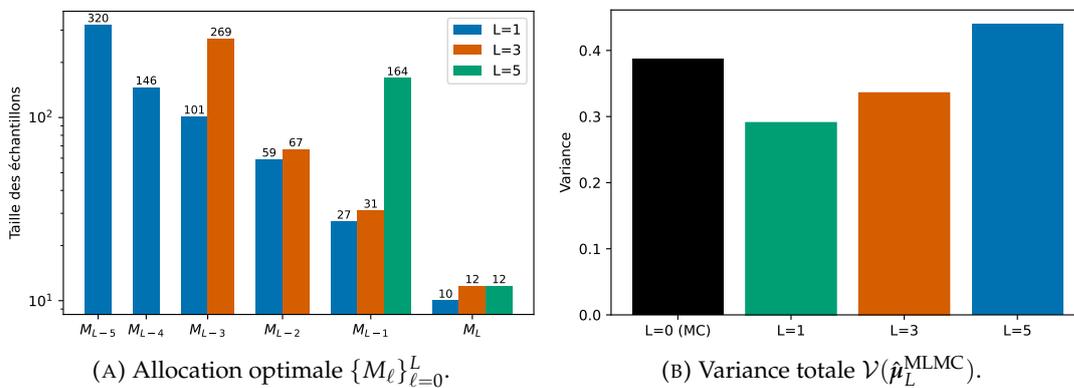


FIGURE 3.14 – Allocation optimale des échantillons sur les différents niveaux et la variance totale résultante des estimateurs MC ( $L = 0$ ) et de différents estimateurs MLMC ( $L \in \{1, 3, 5\}$ ) pour le problème d’estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

les performances du MLMC. Les variances spectrales et cumulées présentées sur la figure 3.15 permettent de visualiser les contributions des différentes composantes à la variance totale.

Comme précédemment, on observe que la variance associée aux grande échelles

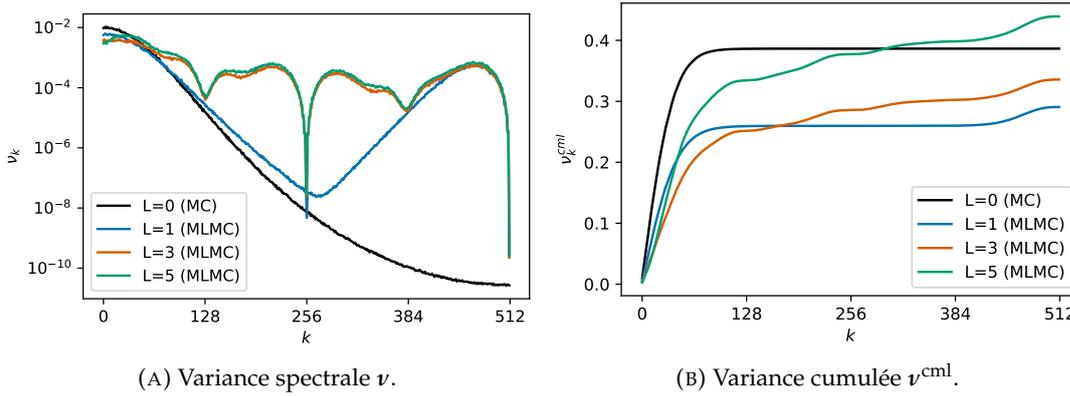


FIGURE 3.15 – Variance spectrale et cumulée de l'estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC ( $L \in \{1, 3, 5\}$ ) pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{001}\mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

est plus faible pour les estimateurs MLMC que pour le MC. Et plus l'estimateur possède de niveaux, plus la variance associée à ces échelles est réduite. Cependant, cette réduction de variance ne concerne que les toutes premières valeurs propres. Pour toutes les autres, le MC a une variance plus faible. Contrairement au cas de l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006}\mathbf{X}_L]$ , les erreurs d'estimation des moyennes et petites échelles ne sont pas négligeables ici. Les contributions des petites échelles sont bien visibles sur la variance cumulative du MLMC à 2 niveaux. Pour le MLMC à 4 et 6 niveaux, les contributions des petites et moyennes échelles sont, en grande partie, celles qui augmentent leur variance totale au delà de celle du MLMC à 2 niveaux.

Pour l'opérateur  $\mathbf{A}_L^{001}$ , dont le spectre est moins concentré sur les grandes échelles, la réduction de variance sur ces grandes échelles ne compense plus forcément les erreurs d'estimation du reste du spectre. Les composantes des petites échelles introduisent trop d'erreur lors de l'application des opérateurs de transfert pour profiter de la réduction de variance d'un estimateur multi-niveaux.

### 3.5 Conclusion

L'utilisation de l'estimateur MLMC pour des champs discrétisés sur des grilles requiert l'utilisation d'opérateurs de transfert (cf. équation (3.7)). Ces opérateurs servent à faire le passage entre les grilles pour permettre des opérations de soustraction entre des champs de même dimension et pour coupler les entrées des modèles. Il a déjà été constaté, dans le contexte similaire des méthodes multigrilles, que de tels opérateurs avaient un impact sur la représentation des différentes échelles d'un champ sur les différentes grilles. Une analyse spectrale avec la base de Hartley (3.21) a permis d'exprimer les effets des opérateurs de transfert choisis (3.19), (3.20) et a montré que ceux-ci introduisent des fines échelles parasites qui polluent la représentation du champ. Ces effets négatifs ont des répercussions sur la variance de l'estimateur MLMC, ce qui a pu être exprimé dans le cas simplifié d'opérateurs linéaires, symétriques et circulants pour un estimateur bi-niveaux. Une illustration 1D respectant les différentes hypothèses faites lors de l'analyse spectrale permet de constater que les effets des opérateurs de transfert conduisent à une mauvaise estimation des moyennes et fines échelles. Les fines échelles parasites causent des erreurs sur la totalité du spectre, ce qui peut contribuer à augmenter substantiellement la

variance totale de l'estimateur. Le traitement de ces erreurs causées par les opérateurs de transfert pourrait donc aboutir à un estimateur avec une variance plus faible, et fait donc l'objet du chapitre suivant.

## Chapitre 4

# Filtered-MLMC

Lors de l'estimation de champs discrétisés, la réduction de variance d'un estimateur MLMC, comparé à un simple estimateur MC, provient d'une meilleure estimation des composantes grandes échelles (basses fréquences). Pour les autres échelles, les effets des opérateurs de transfert de grilles peuvent conduire à une estimation moins bonne que le MC, comme illustré dans la section 3.4. Plus le MLMC comprend de niveaux, plus ces erreurs d'estimation s'étendent à l'ensemble des composantes d'échelles. Dépendamment des simulateurs utilisés, ces erreurs peuvent s'avérer significatives dans la variance totale de l'estimateur. Réduire les impacts négatifs des opérateurs de transfert pourrait permettre, d'abord, d'obtenir de meilleures estimations des moyennes et petites échelles qui ne sont pas bien estimées par le MLMC, et ensuite de réduire la variance totale de l'estimateur. Ce chapitre introduit l'estimateur filtered-MLMC (F-MLMC) qui vise, grâce à l'ajout de filtres passe-bas, à réduire les effets négatifs des opérateurs de transfert. L'idée d'ajouter des filtres provient des méthodes multigrilles [69, 7, 76] où des filtres sont utilisés pour augmenter l'ordre des opérateurs de transfert. On définit l'ordre d'un opérateur de transfert à l'aide du degré maximal d'un polynôme que l'on peut exactement interpoler grâce à cet opérateur. Plus précisément, on dit qu'un opérateur de transfert est d'ordre  $k$  si et seulement si son application laisse invariants les polynômes de degré  $k - 1$ . Avoir des opérateurs de transfert avec ordre suffisamment élevé est une condition de convergence du multigrille géométrique [33]. Sans aboutir à un tel résultat pour les méthodes MLMC, on s'inspire des méthodes multigrilles en ajoutant des filtres aux opérateurs de transfert utilisés et en étudiant l'impact de ces filtres sur la variance de l'estimateur.

### 4.1 Ajout des filtres

Le filtrage des petites échelles d'un champ discrétisé représenté sur une grille de taille  $n_\ell$  se fait avec un filtre passe-bas  $S_\ell: \mathbb{R}^{n_\ell} \rightarrow \mathbb{R}^{n_\ell}$ . On suppose que, pour un vecteur aléatoire  $\mathbf{X}_\ell \in L^2(\Omega, H_\ell)$ , on a  $S_\ell(\mathbf{X}_\ell) \in L^2(\Omega, H_\ell)$ . On souhaite filtrer les petites échelles, qui ne sont pas représentables sur la grille  $\ell - 1$ , d'un champ discrétisé avant de le restreindre pour éviter le phénomène de repliement du spectre présenté dans l'équation (3.56); c'est l'étape de pré-filtrage. On souhaite également faire une étape de post-filtrage, c'est-à-dire filtrer le champ obtenu après la prolongation afin d'enlever les petites échelles parasites présentées dans l'équation (3.48). On utilise ici le même opérateur pour ces deux étapes car l'objectif est le même, i.e., filtrer les petites échelles, mais ce n'est en aucun cas requis. On définit alors les opérateurs de transfert filtrés

$$\bar{P}_{\ell-1}^\ell := S_\ell \circ P_{\ell-1}^\ell \quad \text{et} \quad \bar{R}_\ell^{\ell-1} := R_\ell^{\ell-1} \circ S_\ell, \quad (4.1)$$

pour  $\ell = 1, \dots, L$ . L'ajout des ces étapes de pre- et post-filtrage peut simplement être vu comme la construction de nouveaux opérateurs de transfert qui, dépendamment des opérateurs utilisés, peuvent être d'ordre supérieur aux opérateurs  $P_{\ell-1}^\ell$  et  $R_{\ell-1}^\ell$ . Les opérateurs de transfert entre la grille fine  $L$  et une grille arbitraire  $\ell$  sont

$$\bar{P}_\ell^L = \bar{P}_{L-1}^L \circ \bar{P}_{L-2}^{L-1} \circ \dots \circ \bar{P}_\ell^{\ell+1} \quad \text{et} \quad \bar{R}_L^\ell := \bar{R}_{\ell+1}^\ell \circ \bar{R}_{\ell+2}^{\ell+1} \circ \dots \circ \bar{R}_L^{L-1}, \quad (4.2)$$

pour  $\ell = 0, \dots, L-1$ . On pose  $\bar{P}_L^L = \bar{R}_L^L = id$ . On peut alors définir les simulateurs qui incluent les étapes de pre- et post-filtrage

$$\bar{f}_\ell := \bar{P}_\ell^L \circ f_\ell \circ \bar{R}_L^\ell, \quad \ell = 0, \dots, L. \quad (4.3)$$

Avec les différentes hypothèses faites sur les opérateurs de transfert et de filtrage, on a bien  $\bar{f}_\ell(\mathbf{X}_L) \in L^2(\Omega, H_L)$  pour  $\mathbf{X}_L \in L^2(\Omega, H_L)$ .

#### 4.1.1 Estimateur F-MLMC

La suite  $\{\bar{f}_\ell\}_{\ell=0}^L$  forme la hiérarchie de modèles qui est utilisée pour l'estimateur F-MLMC. On suppose que le coût d'application des opérateurs de transfert et de filtrage est négligable devant le coût d'application du modèle  $f_\ell$  afin de conserver la hiérarchie de coût et de précision de la famille  $\{f_\ell\}_{\ell=0}^L$ . L'estimateur F-MLMC de  $\mathbb{E}[\mathbf{Y}_L]$  s'écrit alors

$$\hat{\mu}_L^{\text{F-MLMC}} = \frac{1}{M_0} \sum_{i=1}^{M_0} \bar{f}_0(\mathbf{X}_L^{(0,i)}) + \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \bar{f}_\ell(\mathbf{X}_L^{(\ell,i)}) - \bar{f}_{\ell-1}(\mathbf{X}_L^{(\ell,i)}), \quad (4.4)$$

avec  $L+1$   $M_\ell$ -échantillons indépendants  $\{\mathbf{X}_L^{(i)}\}_{i=1}^{M_\ell}$  de  $\mathbf{X}_L$ . Cet estimateur F-MLMC est défini de la même manière que l'estimateur MLMC de l'équation (3.7) avec, comme seul changement, l'utilisation de la famille de modèles  $\{\bar{f}_\ell\}_{\ell=0}^L$  plutôt que la famille  $\{f_\ell\}_{\ell=0}^L$  définie dans l'équation (3.6). L'estimateur F-MLMC de l'équation (4.4) étant de la même forme que l'estimateur MLMC du chapitre précédent, son erreur quadratique moyenne est exprimée de la même manière que dans l'équation (3.8).

#### 4.1.2 Filtrés de Shapiro

On choisit de considérer des opérateurs de filtrage linéaires, on identifie donc  $S_\ell$  avec la matrice  $\mathbf{S}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$ . Cela permet d'exprimer plus simplement la variance de l'estimateur F-MLMC et d'avoir des opérateurs dont le coût d'application est faible. On rappelle que l'on a supposé que leur coût d'application était négligable devant celui de l'application des simulateurs. Pour les différents développements qui suivent, on choisit spécifiquement d'utiliser des filtres de Shapiro d'ordre 2 [20, 21, 65]. On définit le filtre de Shapiro d'ordre 2 comme

$$\mathbf{S}_\ell := \frac{1}{4} \begin{bmatrix} 2 & 1 & & & 1 \\ 1 & 2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & 2 & 1 \\ 1 & & & & 1 & 2 \end{bmatrix} \in \mathbb{R}^{n_\ell \times n_\ell}. \quad (4.5)$$

Cette définition du filtre de Shapiro d'ordre 2 est adaptée à des champs discrétisés sur un domaine spatial périodique, ce qui sera le cas de l'illustration de la section 4.3

et du problème du chapitre 5. Un tel filtre a l'avantage d'être une matrice creuse, symétrique et circulante. L'application d'un tel filtre est donc très peu coûteuse et il est possible de le diagonaliser dans la base de Hartley pour étudier son spectre. Le spectre de  $\mathbf{S}_L$  pour un niveau  $L$  associé à une grille de taille 512 est présenté sur la figure 4.1.

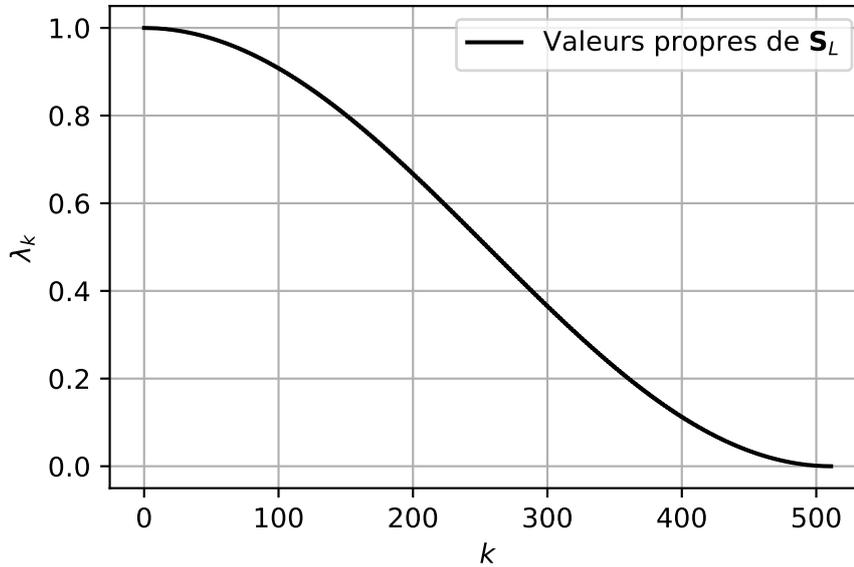


FIGURE 4.1 – Les valeurs propres de la matrice  $\mathbf{S}_L \in \mathbb{R}^{512 \times 512}$  dans la base de Hartley réordonnée  $\mathbf{H}_L^*$ .

Ce filtre est bien un passe-bas, les basses fréquences (ou grandes échelles) sont moins réduites que les hautes fréquences. Il faut cependant noter que toutes les composantes, excepté la première, sont au moins un peu atténuée. Un autre choix d'opérateur de filtrage pourrait permettre de réduire davantage les petites échelles tout en conservant mieux les grandes échelles, mais ce sujet sera abordé dans la section 4.4.

## 4.2 Analyse spectrale

On veut, dans cette section, exprimer les effets de l'ajout des filtres sur la variance de l'estimateur F-MLMC, ce qui permettra de comparer la variance des estimateurs F-MLMC et MLMC. Pour cela, on reprend le même procédé d'analyse spectrale de la section 3.3 que l'on applique ici à l'estimateur F-MLMC. On reprend donc la discrétisation en grilles centrées sur les cellules décrite dans la section 3.2.2, les opérateurs de transfert  $\mathbf{P}_\ell^L$  et  $\mathbf{R}_L^\ell$  des équations (3.19) et (3.20) et on reprend la matrice de Hartley (3.21) pour décomposer un champ discrétisé en ses différentes composantes dans l'espace spectral.

### 4.2.1 Effets de l'opérateur de filtrage sur la base de Hartley

On commence par exprimer les effets des filtres de Shapiro d'ordre 2 sur les matrices de Hartley (3.21). On pose  $\mathbf{S} := \mathbf{S}_\ell$  un filtre de Shapiro d'ordre 2 défini avec l'équation (4.5), où  $\ell$  est fixé entre 1 et  $L$ , et on considère, pour  $k = 0, \dots, n_\ell - 1$ , les

colonnes  $\mathbf{h}_k^\ell$  de la matrice de Hartley  $\mathbf{H}_\ell$ . On a, pour  $j, k = 0, \dots, n_\ell - 1$ ,

$$(\mathbf{S}\mathbf{h}_k^\ell)_j = \frac{1}{4}[(\mathbf{h}_k^\ell)_{j-1} + 2(\mathbf{h}_k^\ell)_j + (\mathbf{h}_k^\ell)_{j+1}]. \quad (4.6)$$

On réutilise la notation  $c_k = \cos(k\pi/n_\ell)$  et on utilise les relations trigonométriques suivantes

$$\cos \frac{(2j-1)k\pi}{n_\ell} + \cos \frac{(2j+1)k\pi}{n_\ell} = 2c_k \cos \frac{2jk\pi}{n_\ell}, \quad (4.7)$$

$$\cos \frac{(2j+1)k\pi}{n_\ell} + \cos \frac{(2j+3)k\pi}{n_\ell} = 2c_k \cos \frac{(2j+2)k\pi}{n_\ell}, \quad (4.8)$$

$$\sin \frac{(2j-1)k\pi}{n_\ell} + \sin \frac{(2j+1)k\pi}{n_\ell} = 2c_k \sin \frac{2jk\pi}{n_\ell}, \quad (4.9)$$

$$\sin \frac{(2j+1)k\pi}{n_\ell} + \sin \frac{(2j+3)k\pi}{n_\ell} = 2c_k \sin \frac{(2j+2)k\pi}{n_\ell}, \quad (4.10)$$

qui permettent d'écrire

$$(\mathbf{S}\mathbf{h}_k^\ell)_j = \frac{c_k}{2\sqrt{n_\ell}} \left[ \cos \frac{2jk\pi}{n_\ell} + \cos \frac{(2j+2)k\pi}{n_\ell} + \sin \frac{2jk\pi}{n_\ell} + \sin \frac{(2j+2)k\pi}{n_\ell} \right] \quad (4.11)$$

$$= \frac{c_k}{2\sqrt{n_\ell}} \left[ 2c_k \cos \frac{(2j+1)k\pi}{n_\ell} + 2c_k \sin \frac{(2j+1)k\pi}{n_\ell} \right] \quad (4.12)$$

$$= c_k^2 (\mathbf{h}_k^\ell)_j. \quad (4.13)$$

Le filtre de Shapiro d'ordre 2 atténue donc les différents colonnes de la base de Hartley par un facteur  $c_k^2$  compris entre 0 et 1. Ce facteur est proche de 1 pour les composantes grandes échelles (quand  $k$  tend vers 0 ou vers  $n_\ell - 1$ ) et proche de 0 pour les composantes petites échelles (quand  $k$  tend vers  $n_\ell/2$ ). Les coefficients  $c_k^2$  sont les valeurs propres de  $\mathbf{S}$  car on a  $\mathbf{S}\mathbf{H}_L = \mathbf{H}_L\mathbf{\Lambda}$  où  $\mathbf{\Lambda} = \text{Diag}(\{c_k^2\}_{k=0}^{n_\ell-1})$ . Elles sont donc visualisables sur la figure 4.1 où elles sont ordonnées de manière décroissante.

À partir de l'équation (3.57), on peut facilement déduire les relations

$$\bar{\mathbf{P}}\mathbf{H}_{\ell-1} = \sqrt{2}\mathbf{H}_\ell\mathbf{C}_3^T, \quad \bar{\mathbf{R}}\mathbf{H}_\ell = \mathbf{H}_{\ell-1}\mathbf{C}_3, \quad (4.14)$$

où  $\mathbf{C}_3 := [\text{Diag}(\{c_k^3\}_{k=0}^{n_\ell-1}) \text{Diag}(\{-c_{n_\ell-1+k}^3\}_{k=0}^{n_\ell-1})] \in \mathbb{R}^{n_{\ell-1} \times n_\ell}$ . À partir de l'équation (3.58), on obtient les relations

$$\mathbf{H}_\ell^T \bar{\mathbf{P}} = \sqrt{2}\mathbf{C}_3^T \mathbf{H}_{\ell-1}^T, \quad \mathbf{H}_{\ell-1}^T \bar{\mathbf{R}} = \mathbf{C}_3 \mathbf{H}_\ell^T. \quad (4.15)$$

## 4.2.2 Impact sur la variance du F-MLMC

### Modèles linéaires

On suppose que les simulateurs  $f_\ell$  sont linéaires pour  $\ell = 0, \dots, L$ . On les identifie avec les matrices  $\mathbf{F}_\ell \in \mathbb{R}^{n_\ell \times n_\ell}$  et on définit

$$\bar{\mathbf{F}}_\ell = \bar{\mathbf{P}}_\ell^L \mathbf{F}_\ell \bar{\mathbf{R}}_\ell^L, \quad (4.16)$$

avec  $\bar{\mathbf{P}}_L^L = \bar{\mathbf{R}}_L^L = \mathbf{I}_{n_L}$ . Ces matrices forment la hiérarchie de simulateurs utilisée par le F-MLMC. La variance de l'estimateur F-MLMC peut alors s'écrire, de manière

similaire à l'équation (3.70),

$$\mathcal{V}(\hat{\mu}_L^{\text{F-MLMC}}) = \frac{1}{M_0} \|\bar{\mathbf{F}}_0 \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2 + \sum_{\ell=1}^L \frac{1}{M_\ell} \|(\bar{\mathbf{F}}_\ell - \bar{\mathbf{F}}_{\ell-1}) \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2, \quad (4.17)$$

avec  $\mathbf{G} := \mathbb{E}[\dot{\mathbf{X}}_L \dot{\mathbf{X}}_L^T]$  la matrice de covariance de  $\dot{\mathbf{X}}_L := \mathbf{X}_L - \mathbb{E}[\mathbf{X}_L]$  admettant la décomposition  $\mathbf{G} = \mathbf{G}^{1/2} \mathbf{G}^{1/2}$ . On a la décomposition dans la base de Hartley

$$\mathcal{V}(\hat{\mu}_L^{\text{F-MLMC}}) = \frac{1}{M_0} \|\mathbf{H}_L^T \bar{\mathbf{F}}_0 \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2 + \sum_{\ell=1}^L \frac{1}{M_\ell} \|\mathbf{H}_L^T (\bar{\mathbf{F}}_\ell - \bar{\mathbf{F}}_{\ell-1}) \mathbf{G}^{1/2}\|_{F, \mathbf{W}_L}^2. \quad (4.18)$$

On retrouve exactement la relation (3.80) obtenue pour le MLMC, mais avec les nouveaux simulateurs  $\bar{\mathbf{F}}_\ell$ . La question est alors de savoir si l'ajout des filtres permet de réduire la norme de la différence entre deux simulateurs définis sur des niveaux successifs, i.e., est-ce que  $\|\bar{\mathbf{F}}_\ell - \bar{\mathbf{F}}_{\ell-1}\|_{F, \mathbf{W}_L}$  est plus petit que  $\|\bar{\mathbf{F}}_\ell - \bar{\mathbf{F}}_{\ell-1}\|_{F, \mathbf{W}_L}$ .

### Modèles linéaires, symétriques circulants

Pour exprimer la différence entre deux simulateurs, on reprend l'hypothèse des opérateurs  $\mathbf{F}_\ell$  symétriques et circulants pour  $\ell = 0, \dots, L$  et on se restreint à l'étude d'un estimateur à bi-niveaux. On a donc, pour  $\ell \in \{0, 1\}$ , les deux simulateurs linéaires, symétrique et circulant,  $\mathbf{F}_0$  et  $\mathbf{F}_1$ , qui sont tels que

$$\mathbf{F}_\ell = \mathbf{H}_\ell \boldsymbol{\Lambda}_\ell \mathbf{H}_\ell^T, \quad \boldsymbol{\Lambda}_\ell := \text{Diag}(\{\lambda_k^\ell\}_{k=0}^{n_\ell-1}). \quad (4.19)$$

On note  $\bar{\mathbf{P}} = \mathbf{S}_1 \mathbf{P}_0^1$  et  $\bar{\mathbf{R}} = \mathbf{R}_1^0 \mathbf{S}_1$  les opérateurs de transfert filtrés entre le niveau fin  $\ell = 1$  et le niveau grossier  $\ell = 0$ . En partant de la définition de  $\bar{\mathbf{F}}_0$  de l'équation (4.16) et en utilisant les relations des équations (4.14) et (4.15), on obtient

$$\bar{\mathbf{F}}_0 = \bar{\mathbf{P}} \mathbf{F}_0 \bar{\mathbf{R}} = \bar{\mathbf{P}} \mathbf{H}_0 \boldsymbol{\Lambda}_0 \mathbf{H}_0^T \bar{\mathbf{R}} = \sqrt{2} \mathbf{H}_1 \mathbf{C}_3^T \boldsymbol{\Lambda}_0 \mathbf{C}_3 \mathbf{H}_1^T. \quad (4.20)$$

On définit alors

$$\bar{\mathbf{M}} = \sqrt{2} \mathbf{C}_3^T \boldsymbol{\Lambda}_0 \mathbf{C}_3 = \begin{bmatrix} \bar{\mathbf{M}}_{11} & \bar{\mathbf{M}}_{12} \\ \bar{\mathbf{M}}_{21} & \bar{\mathbf{M}}_{22} \end{bmatrix}, \quad (4.21)$$

avec

$$\bar{\mathbf{M}}_{11} := \sqrt{2} \text{Diag}(\{c_k^6 \lambda_k^0\}_{k=0}^{n_0-1}), \quad (4.22)$$

$$\bar{\mathbf{M}}_{22} := \sqrt{2} \text{Diag}(\{c_{n_0+k}^6 \lambda_k^0\}_{k=0}^{n_0-1}), \quad (4.23)$$

$$\bar{\mathbf{M}}_{12} := \sqrt{2} \text{Diag}(\{-c_k^3 c_{n_0+k}^3 \lambda_k^0\}_{k=0}^{n_0-1}) = \bar{\mathbf{M}}_{21}. \quad (4.24)$$

L'équation (4.17), pour l'estimateur F-MLMC bi-niveaux, devient

$$\mathcal{V}(\hat{\mu}_1^{\text{F-MLMC}}) = \frac{1}{M_0} \|\bar{\mathbf{M}} \mathbf{H}_1^T \mathbf{G}^{1/2}\|_{F, \mathbf{W}_1}^2 + \frac{1}{M_1} \|(\boldsymbol{\Lambda}_1 - \bar{\mathbf{M}}) \mathbf{H}_1^T \mathbf{G}^{1/2}\|_{F, \mathbf{W}_1}^2. \quad (4.25)$$

On retrouve une expression très similaire à celle de la variance de l'estimateur MLMC bi-niveaux dans le cas de modèles linéaires, symétriques et circulants de l'équation (3.87). La variance du MLMC fait apparaître la matrice  $\mathbf{M}$ , définie dans l'équation (3.83), tandis que la variance du F-MLMC fait apparaître la matrice  $\bar{\mathbf{M}}$ . Cette matrice  $\bar{\mathbf{M}}$  diffère de la matrice  $\mathbf{M}$  par la puissance des facteurs  $c_k$  qui est multipliée par 3. Les facteurs  $c_k$  étant compris entre 1 et -1, l'augmentation de la puissance réduit

leur valeur et les fait tendre plus vite vers 0. Les blocs extra-diagonaux  $\bar{\mathbf{M}}_{12}$  et  $\bar{\mathbf{M}}_{21}$  ont alors une norme de Frobenius plus faible que les blocs  $\mathbf{M}_{12}$  et  $\mathbf{M}_{21}$ . La réduction de la norme de ces blocs contribue à réduire la norme de la différence  $\Lambda_1 - \bar{\mathbf{M}}$ . Pour étudier les changements sur les blocs diagonaux  $\bar{\mathbf{M}}_{11}$  et  $\bar{\mathbf{M}}_{22}$ , on aimerait pouvoir exprimer les valeurs propres de  $\mathbf{F}_0$  en fonction de celles de  $\mathbf{F}_1$ .

### Opérateurs de Galerkin

Pour exprimer les valeurs propres de  $\mathbf{F}_0$  en fonction de celles de  $\mathbf{F}_1$ , on souhaite utiliser l'opérateur de Galerkin. On peut réutiliser l'opérateur défini dans l'équation (3.88) avec les opérateurs de prolongation et de restriction  $\mathbf{P}$  et  $\mathbf{R}$ , ou alors définir un opérateur basé sur  $\bar{\mathbf{P}}$  et  $\bar{\mathbf{R}}$ . On considère donc

$$\mathbf{F}_0 := \frac{1}{2}\mathbf{R}\mathbf{F}_1\mathbf{P}, \quad \text{et} \quad \mathbf{F}_0^* := \frac{1}{2}\bar{\mathbf{R}}\mathbf{F}_1\bar{\mathbf{P}}, \quad (4.26)$$

où  $\mathbf{F}_0$  est exactement le même opérateur que celui utilisé dans la section 3.3.3. Il est important de noter que la proposition 1, portant sur l'optimalité de l'opérateur de Galerkin, n'est plus vérifiée si on utilise les opérateurs filtrés  $\bar{\mathbf{P}}$  et  $\bar{\mathbf{R}}$ . En effet, la preuve de cette propriété repose sur la relation  $\mathbf{R}\mathbf{R}^\top = \mathbf{I}$  qui n'est pas vérifiée pour  $\bar{\mathbf{R}}$  car

$$\bar{\mathbf{R}}\bar{\mathbf{R}}^\top = \mathbf{R}\mathbf{S}(\mathbf{R}\mathbf{S})^\top = \mathbf{R}\mathbf{S}^2\mathbf{R}^\top, \quad (4.27)$$

et  $\mathbf{S}^2 \neq \mathbf{I}$ . L'ajout des filtres ne permet donc plus de garantir que l'opérateur  $\mathbf{F}_0$ , ou  $\mathbf{F}_0^*$ , soient solutions de  $\arg \min_{\mathbf{A}_0} \|\mathbf{F}_1 - \bar{\mathbf{P}}\mathbf{A}_0\bar{\mathbf{R}}\|_{F, \mathbf{W}_1}^2$ .

Sans connaissance de l'opérateur optimal à utiliser, on présente les résultats pour  $\mathbf{F}_0$  et  $\mathbf{F}_0^*$ . L'opérateur  $\mathbf{F}_0$  est diagonalisable dans la base de Hartley (cf. l'équation (3.100)) et on avait  $\Lambda_0 = \sqrt{2}^{-1}\mathbf{C}\Lambda_1\mathbf{C}^\top$  et

$$\lambda_k^0 = \frac{1}{\sqrt{2}}(c_k^2\lambda_k^1 + c_{n_0+k}^2\lambda_{n_0+k}^1). \quad (4.28)$$

Pour  $\mathbf{F}_0^*$ , on a

$$\mathbf{F}_0^* = \frac{1}{2}\bar{\mathbf{R}}\mathbf{F}_1\bar{\mathbf{P}} = \frac{1}{2}\bar{\mathbf{R}}\mathbf{H}_1\Lambda_1\mathbf{H}_1^\top\bar{\mathbf{P}} = \frac{\sqrt{2}}{2}\mathbf{H}_1\mathbf{C}_3\Lambda_1\mathbf{C}_3^\top\mathbf{H}_1^\top. \quad (4.29)$$

On simplifie les expressions suivantes en ommettant les indices de début et de fin des suites qui vont, ici, toujours de  $k = 0$  à  $k = n_0 - 1$ . Ainsi,  $\{c_k\} = \{c_k\}_{k=0}^{n_0-1}$ , et on simplifie de la même manière l'écriture de toutes les autres suites. On a

$$\mathbf{C}_3\Lambda_1\mathbf{C}_3^\top = \begin{bmatrix} \text{Diag}(\{c_k^3\}) \\ \text{Diag}(\{-c_{n_0+k}^3\}) \end{bmatrix}^\top \begin{bmatrix} \text{Diag}(\{\lambda_k^1\}) & 0 \\ 0 & \text{Diag}(\{\lambda_{n_0+k}^1\}) \end{bmatrix} \begin{bmatrix} \text{Diag}(\{c_k^3\}) \\ \text{Diag}(\{-c_{n_0+k}^3\}) \end{bmatrix} \quad (4.30)$$

$$= \text{Diag}(\{c_k^6\lambda_k^1\}) + \text{Diag}(\{c_{n_0+k}^6\lambda_{n_0+k}^1\}). \quad (4.31)$$

La matrice  $\Lambda_0^* := \sqrt{2}^{-1}\mathbf{C}_3\Lambda_1\mathbf{C}_3^\top$  est donc la matrice diagonale des valeurs propres de  $\mathbf{F}_0^*$  et on peut exprimer les valeurs propres  $\{(\lambda_k^0)^*\}_{k=0}^{n_0-1}$  de  $\mathbf{F}_0^*$  en fonction de celles de  $\mathbf{F}_1$ ,

$$(\lambda_k^0)^* = \frac{1}{\sqrt{2}}(c_k^6\lambda_k^1 + c_{n_0+k}^6\lambda_{n_0+k}^1), \quad \forall k = 0, \dots, n_0 - 1. \quad (4.32)$$

Les blocs de la matrice  $\bar{\mathbf{M}}$ , définie comme (4.21) dans la section précédente, peuvent maintenant s'écrire

$$\bar{\mathbf{M}}_{11} := \text{Diag}(\{c_k^8 \lambda_k^1 + c_k^6 c_{n_0+k}^2 \lambda_{n_0+k}^1\}_{k=0}^{n_0-1}), \quad (4.33)$$

$$\bar{\mathbf{M}}_{22} := \text{Diag}(\{c_{n_0+k}^8 \lambda_k^1 + c_k^2 c_{n_0+k}^6 \lambda_k^1\}_{k=0}^{n_0-1}) = \text{Diag}(\{c_k^8 \lambda_k^1 + c_{k-n_0}^2 c_k^6 \lambda_{k-n_0}^1\}_{k=n_0}^{2n_0-1}), \quad (4.34)$$

$$\bar{\mathbf{M}}_{12} := \text{Diag}(\{-c_k^5 c_{n_0+k}^3 \lambda_k^1 - c_k^3 c_{n_0+k}^5 \lambda_{n_0+k}^1\}_{k=0}^{n_0-1}) = \bar{\mathbf{M}}_{21}. \quad (4.35)$$

On définit également la matrice

$$\bar{\mathbf{M}}^* := \sqrt{2} \mathbf{C}_3^T \Lambda_0^* \mathbf{C}^3, \quad (4.36)$$

dont les blocs s'écrivent

$$\bar{\mathbf{M}}_{11}^* := \text{Diag}(\{c_k^{12} \lambda_k^1 + c_k^6 c_{n_0+k}^6 \lambda_{n_0+k}^1\}_{k=0}^{n_0-1}), \quad (4.37)$$

$$\bar{\mathbf{M}}_{22}^* := \text{Diag}(\{c_{n_0+k}^{12} \lambda_k^1 + c_k^6 c_{n_0+k}^6 \lambda_k^1\}_{k=0}^{n_0-1}) = \text{Diag}(\{c_k^{12} \lambda_k^1 + c_{k-n_0}^6 c_k^6 \lambda_{k-n_0}^1\}_{k=n_0}^{2n_0-1}), \quad (4.38)$$

$$\bar{\mathbf{M}}_{12}^* := \text{Diag}(\{-c_k^9 c_{n_0+k}^3 \lambda_k^1 - c_k^3 c_{n_0+k}^9 \lambda_{n_0+k}^1\}_{k=0}^{n_0-1}) = \bar{\mathbf{M}}_{21}^*. \quad (4.39)$$

On peut alors comparer les expressions des blocs des matrices  $\bar{\mathbf{M}}$  et  $\bar{\mathbf{M}}^*$  avec ceux de la matrice  $\mathbf{M}$ , donnés dans les équations (3.104) à (3.106), et qui apparaissent dans l'expression de la variance du MLMC. On rappelle que, au vu des expressions des variances du MLMC (3.87) et du F-MLMC (4.25), on cherche à trouver le cas où la norme de la différence entre  $\Lambda_1$  et  $\mathbf{M}$  est la plus faible possible. On observe, comme précédemment, que les expressions des blocs de  $\bar{\mathbf{M}}$  et  $\bar{\mathbf{M}}^*$  sont très similaires à celles des blocs de  $\mathbf{M}$ , mais avec une augmentation des puissances sur les coefficients  $c_k$ . Ces puissances sont plus élevées pour  $\bar{\mathbf{M}}^*$  qui contient 4 applications des opérateurs de filtrage  $\mathbf{S}$  contre 2 applications pour  $\bar{\mathbf{M}}$ . La diagonale de  $\bar{\mathbf{M}}$  (resp.  $\bar{\mathbf{M}}^*$ ) diffère de  $\Lambda_1$  par les facteurs réducteurs  $c_k^8$  (resp.  $c_k^{12}$ ) devant les valeurs propres  $\lambda_k^1$ , et par les termes parasites des valeurs propres complémentaires  $\lambda_{n_0+k}^1$ , heureusement réduits par un facteur  $c_k^6 c_{n_0+k}^2$  (resp.  $c_k^6 c_{n_0+k}^6$ ). La figure 4.2 représente les coefficients des matrices  $\mathbf{M}$ ,  $\bar{\mathbf{M}}$  et  $\bar{\mathbf{M}}^*$ . La sous-figure de gauche est donc exactement la figure 3.7 du chapitre précédent, remise ici pour la comparaison.

L'ajout des filtres de Shapiro d'ordre 2 réduit considérablement les facteurs associés aux valeurs propres parasites sur la diagonale (courbe bleue) ainsi que les facteurs des termes extra-diagonaux (courbe orange). Plus l'opérateur  $\mathbf{S}$  est appliqué, plus la réduction est importante, comme on peut le voir en comparant les courbes de  $\bar{\mathbf{M}}$  et  $\bar{\mathbf{M}}^*$ . Ces changements contribuent à réduire la norme de la différence  $\Lambda_1 - \bar{\mathbf{M}}$  (ou  $\Lambda_1 - \bar{\mathbf{M}}^*$ ) et démontrent l'efficacité du filtre pour atténuer les effets négatifs des opérateurs de transfert. Cependant, on peut également observer que les facteurs associés aux valeurs propres correctes de l'opérateur  $\mathbf{F}_1$ , que l'on souhaite approximer le mieux possible, sont également réduites. La moitié du spectre associée aux composantes petites échelles,  $n_1/4 \leq k < 3n_1/4$ , est quasiment nulle, tandis que les valeurs propres associées aux grandes échelles sont fortement atténuées quand on les compare aux facteurs de la matrice  $\mathbf{M}$ . Ce changement contribue à augmenter la norme de la différence  $\Lambda_1 - \bar{\mathbf{M}}$  (ou  $\Lambda_1 - \bar{\mathbf{M}}^*$ ). Ce résultat était attendu, l'ajout d'une étape pour filtrer les petites échelles ne permet plus de les représenter avec le simulateur basse fidélité  $\mathbf{F}_0$ . Pour que l'étape de filtrage soit bénéfique, il faut s'assurer d'un bon compromis entre la réduction des effets négatifs introduits par les opérateurs de transfert et la perte d'information du modèle haute fidélité  $\mathbf{F}_1$ . On a, ici, un exemple

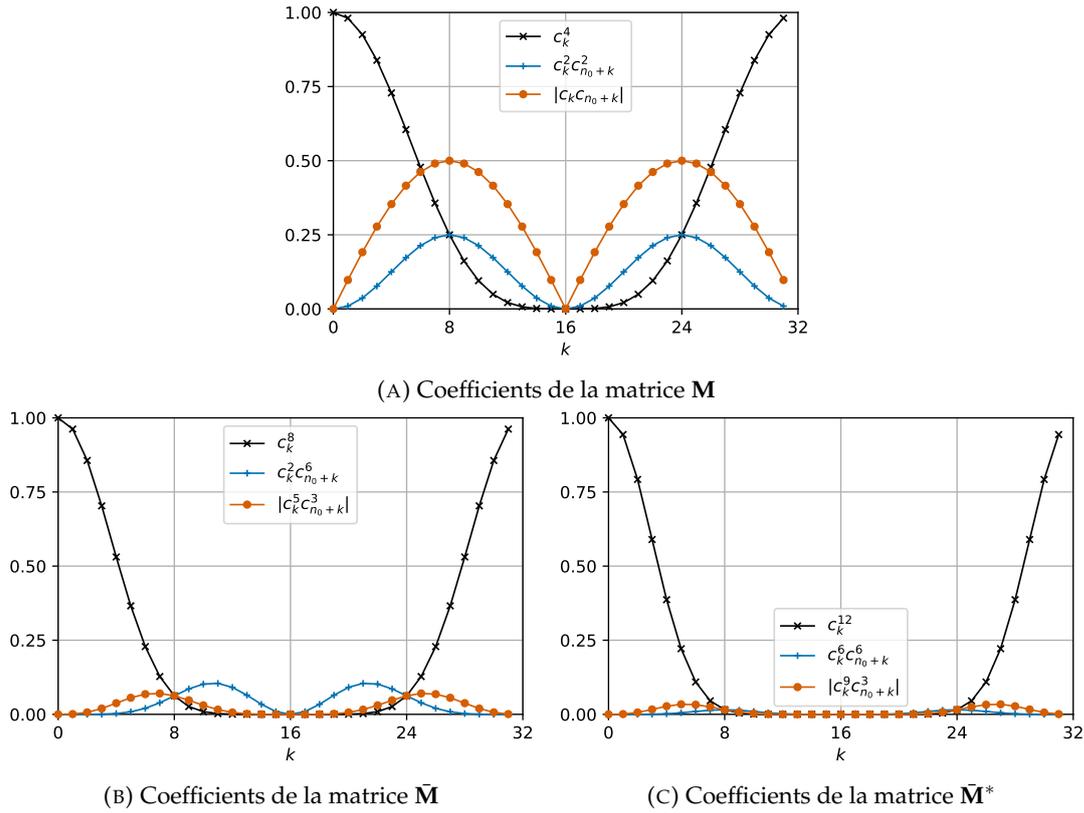


FIGURE 4.2 – Les facteurs devant les valeurs propres dans les matrices  $\mathbf{M}$  (A),  $\tilde{\mathbf{M}}$  (B) et  $\tilde{\mathbf{M}}^*$  (C) en fonction de  $k = 0, \dots, n_1 - 1$ . La courbe noire correspond aux facteurs devant la bonne valeur propre sur la diagonale, tandis que la courbe bleue est associée aux facteurs devant le terme parasite des valeurs propres complémentaires. Finalement la courbe orange correspond aux facteurs des blocs extra-diagonaux en valeur absolue.

d'un tel compromis avec la matrice  $\tilde{\mathbf{M}}^*$  qui réduit plus fortement les termes parasites comparé à  $\tilde{\mathbf{M}}$ , mais au prix d'une perte d'information plus importante sur les valeurs propres associées aux grandes échelles de  $\mathbf{F}_1$ .

### 4.3 Illustration 1D

On souhaite maintenant tester l'estimateur F-MLMC et le comparer au MLMC. L'analyse spectrale a montré que, dans le cas spécifique d'un simulateur linéaire, symétrique et circulant, les filtres contribuent à réduire les effets négatifs des opérateurs de transfert au prix d'une perte d'information sur le simulateur utilisé. En fonction du problème considéré, les filtres peuvent réduire la variance totale de l'estimateur comparé au MLMC. Tester l'estimateur F-MLMC sur une illustration simple permet de constater l'efficacité de l'ajout des filtres. Pour cela, on reprend l'illustration présentée dans la section 3.4. On cherche donc à estimer  $\mathbb{E}[\mathbf{A}_L \mathbf{X}_L] = \mathbf{0}_{n_L}$  où  $\mathbf{X}_L \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{I}_{n_L})$  et  $\mathbf{A}_L$  est une matrice symétrique et circulante. On utilisera les matrices  $\mathbf{A}_L^{006}$  et  $\mathbf{A}_L^{001}$  données dans la figure 3.8 et dont les valeurs propres sont présentées dans la figure 3.9. On reprend la même hiérarchie de grilles en partant d'une grille fine dont la taille est fixée à 512, et le même budget total  $\eta = 100C_L$  pour tous les estimateurs. Pour l'estimateur MLMC, on reprend les simulateurs utilisés

dans l'illustration du chapitre précédent. On teste ici deux estimateurs F-MLMC : le premier utilise les simulateurs basés sur les opérateurs de Galerkin non-filtrés

$$\bar{f}_\ell = \bar{\mathbf{P}}_\ell^L \mathbf{R}_L^\ell \mathbf{A}_L \mathbf{P}_\ell^L \bar{\mathbf{R}}_L^\ell, \quad (4.40)$$

et le second les simulateurs basés sur les opérateurs de Galerkin filtrés

$$\bar{f}_\ell^* = \bar{\mathbf{P}}_\ell^L \bar{\mathbf{R}}_L^\ell \mathbf{A}_L \bar{\mathbf{P}}_\ell^L \bar{\mathbf{R}}_L^\ell. \quad (4.41)$$

On différencie ces estimateurs en nommant le second F-MLMC\*. La plupart des résultats étant très similaires pour ces deux estimateurs, on se concentre principalement sur l'estimateur F-MLMC. La comparaison entre le F-MLMC et le F-MLMC\* sera présentée à la fin de cette section avec la variance totale des deux estimateurs. On rappelle qu'en pratique les opérateurs de Galerkin sont rarement utilisés car ils nécessitent d'être stockés explicitement. Une alternative est d'utiliser des opérateurs de basse fidélité construits par rediscrétisation de l'opérateur haute fidélité. C'est ce qui est fait dans les applications du chapitre 5 où ces opérateurs sont stockés sous la forme d'une application sur des vecteurs.

Une première visualisation intéressante est celle d'une simple estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  et  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  par un MC, un MLMC ou un F-MLMC. La figure 4.3 présente un exemple de ces estimations en considérant 6 niveaux et avec le budget fixé  $\eta = 1000\mathcal{C}_L$ .

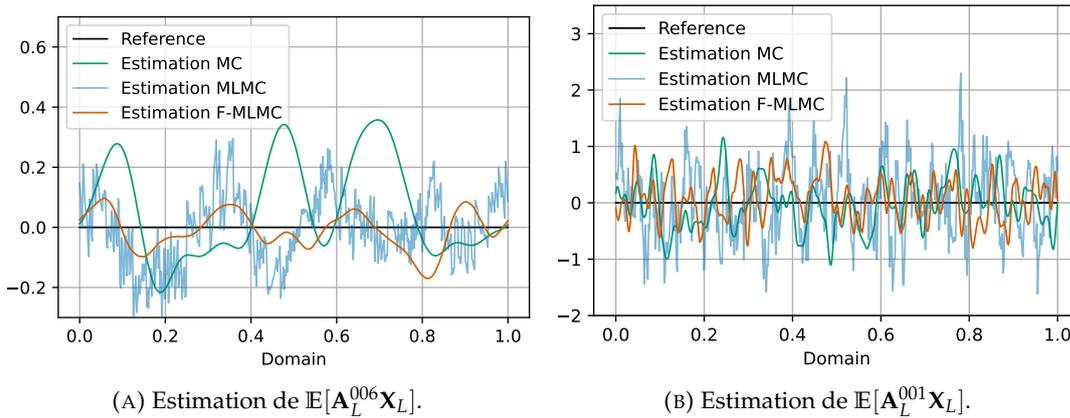


FIGURE 4.3 – Estimation de  $\mathbb{E}[\mathbf{A}_L \mathbf{X}_L]$  par un MC, un MLMC et un F-MLMC à 6 niveaux ( $L = 5$ ) avec le budget total  $\eta = 1000\mathcal{C}_L$  et les allocations optimales présentées dans les figures 4.5a et 4.8a.

On observe bien les différences entre les spectres de  $\mathbf{A}_L^{006}$  et  $\mathbf{A}_L^{001}$ , présentés dans la figure 3.9, sur les estimation par MC. L'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  contient des échelles bien plus fines que celle de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ , ce qui est dû à la longueur de corrélation  $D$  qui est plus faible. Cette différence est cependant moins visible sur les estimations par MLMC. Que ce soit pour estimer  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  ou  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ , le résultat contient de nombreuses composantes petites échelles. Ce sont les composantes parasites introduites par les opérateurs de transfert qui conduisent, comme vu plus haut, à une mauvaise estimation de ces composantes et, potentiellement, à une détérioration de la variance totale de l'estimateur. L'effet de l'ajout des filtres est alors clairement visible dans l'estimation par F-MLMC, les composantes petites échelles (hautes fréquences) ont quasiment disparu. Pour avoir plus d'indication sur la qualité des estimateurs, on regarde d'abord l'estimation des paramètres  $\alpha$  et  $\beta$  du théorème 1 obtenues par regression des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  de l'équation (3.109) pour  $\ell = 0, \dots, L$  et avec  $L = 5$ .

Ces quantités sont présentées sur la figure 4.4 pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ .

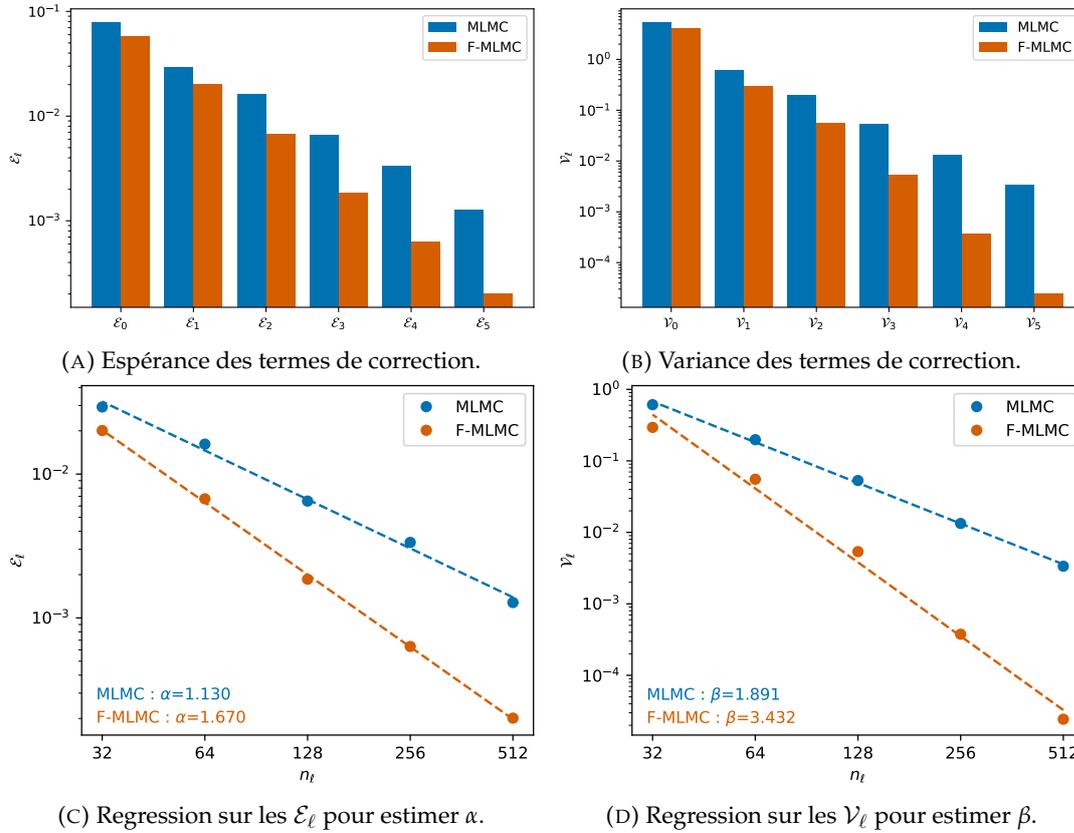


FIGURE 4.4 – Estimation des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  (cf. équation (3.109)) pour  $\ell = 0, \dots, L$  (haut) et regression sur ces quantités afin d'estimer  $\alpha$  et  $\beta$  (bas) pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Les estimations sont faites à partir de 1000 échantillons.

L'estimateur MLMC et F-MLMC semblent, tous deux, respecter les relations, supposées dans le théorème 1,  $\mathcal{E}_\ell \lesssim n_\ell^{-\alpha}$  et  $\mathcal{V}_\ell \lesssim n_\ell^{-\beta}$  pour  $\ell = 0, \dots, L$ . Le F-MLMC conduit à des valeurs de  $\alpha$  et  $\beta$  presque deux fois plus élevées que celles du MLMC, et les  $\mathcal{V}_\ell$  du F-MLMC sont toutes plus faibles que celles du MLMC. On s'attend donc à une réduction de l'erreur quadratique moyenne pour le même budget. C'est effectivement ce que l'on observe sur la figure 4.5 qui présente l'allocation optimale des estimateurs MLMC et F-MLMC à 6 niveaux ainsi que la variance totale des estimateurs pour différentes valeurs de  $L$ .

La variance totale des estimateurs F-MLMC est plus faible que celle des estimateurs MLMC, et ce, peu importe le nombre de niveaux utilisés. Cette réduction de la variance est plus importante quand plus de niveaux sont considérés. Plus il y a de niveaux, plus les effets négatifs des opérateurs de transfert sont importants et donc plus l'ajout des filtres est utile. L'allocation optimale pour les estimateurs à 6 niveaux montre que l'ajout des filtres permet de réduire considérablement la taille des ensembles sur les niveaux fins, laissant alors une grande partie du budget pour augmenter la taille des ensembles sur les grilles grossières. La décomposition spectrale de la variance des différents estimateurs est présentée sur la figure 4.6.

En comparant la décomposition spectrale de la variance des estimateurs MLMC et F-MLMC, on observe une réduction de la variance de toutes les composantes. La réduction de la variance des premières composantes est difficilement visible

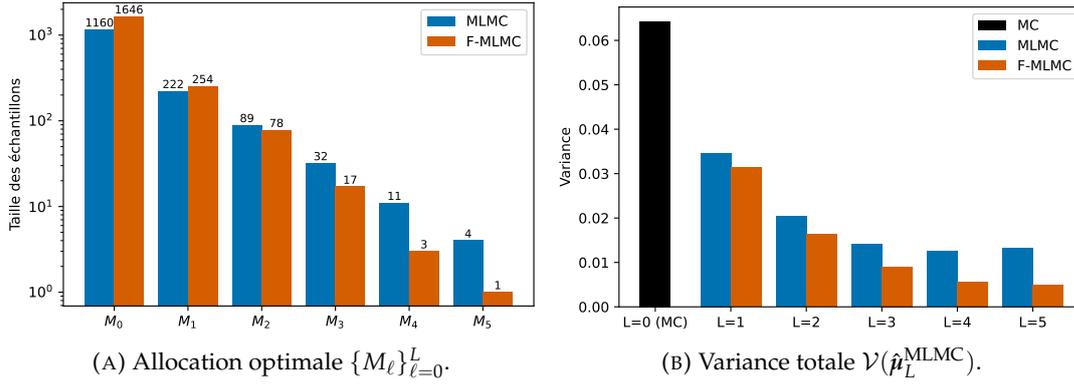


FIGURE 4.5 – Allocation optimale des échantillons pour les estimateurs MLMC et F-MLMC à 6 niveaux, et la variance totale résultante des estimateurs MC ( $L = 0$ ) et de différents estimateurs MLMC et F-MLMC pour le problème d’estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

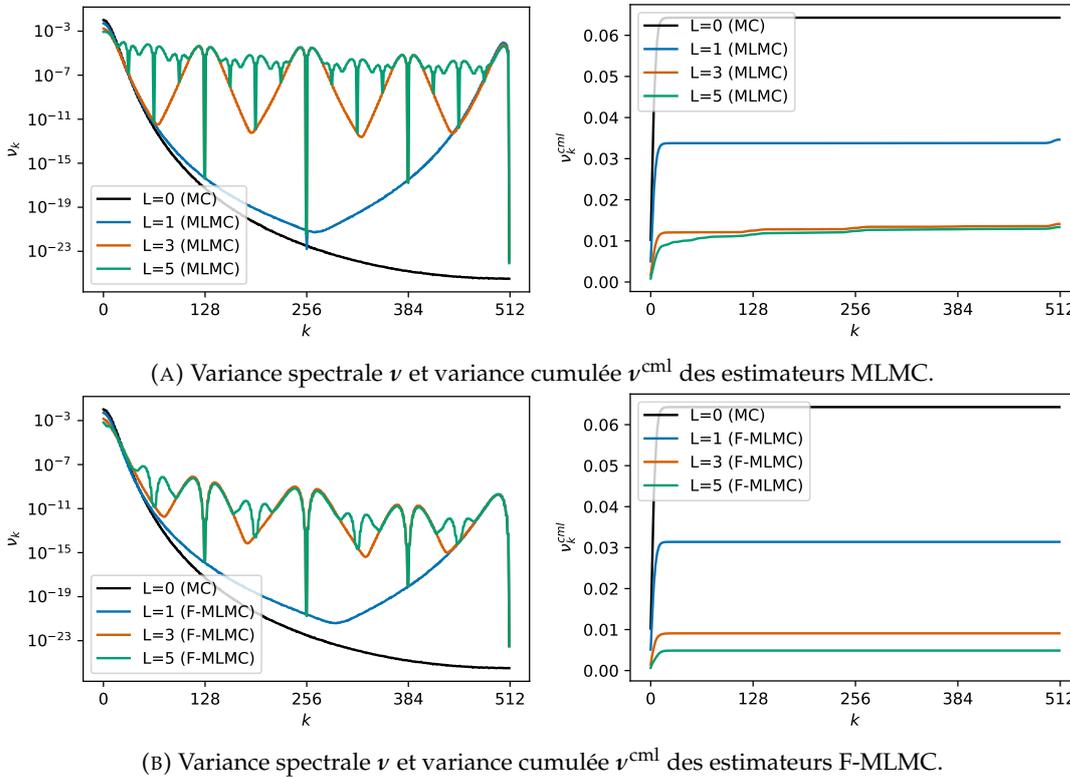


FIGURE 4.6 – Variance spectrale et cumulée de l’estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC (haut) et F-MLMC (bas), avec  $L \in \{1, 3, 5\}$ , pour le problème d’estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

sur les courbes de la variance spectrale  $\nu$  mais elle apparait sur les courbes de la variance cumulée  $\nu^{\text{cml}}$  où la valeur atteinte après la prise en compte des premières composantes est plus faible pour le F-MLMC. Le filtrage des petites échelles conduit à une meilleure estimation de tout le spectre, et donc à une réduction de la variance

totale. Pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006}\mathbf{X}_L]$ , où l'estimateur MLMC est déjà bien meilleur que le MC, l'ajout des filtres permet de réduire encore plus la variance totale de l'estimateur en ayant une meilleure estimation de toutes les composantes.

La même série de résultats est maintenant présentée pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{001}\mathbf{X}_L]$ . La section 3.4 a montré que, pour ce problème, les grilles les plus grossières ne permettent pas une représentation suffisamment fidèle du champ étudié et leur utilisation détériore la variance du MLMC. On commence par regarder les quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  avec la figure 4.7.

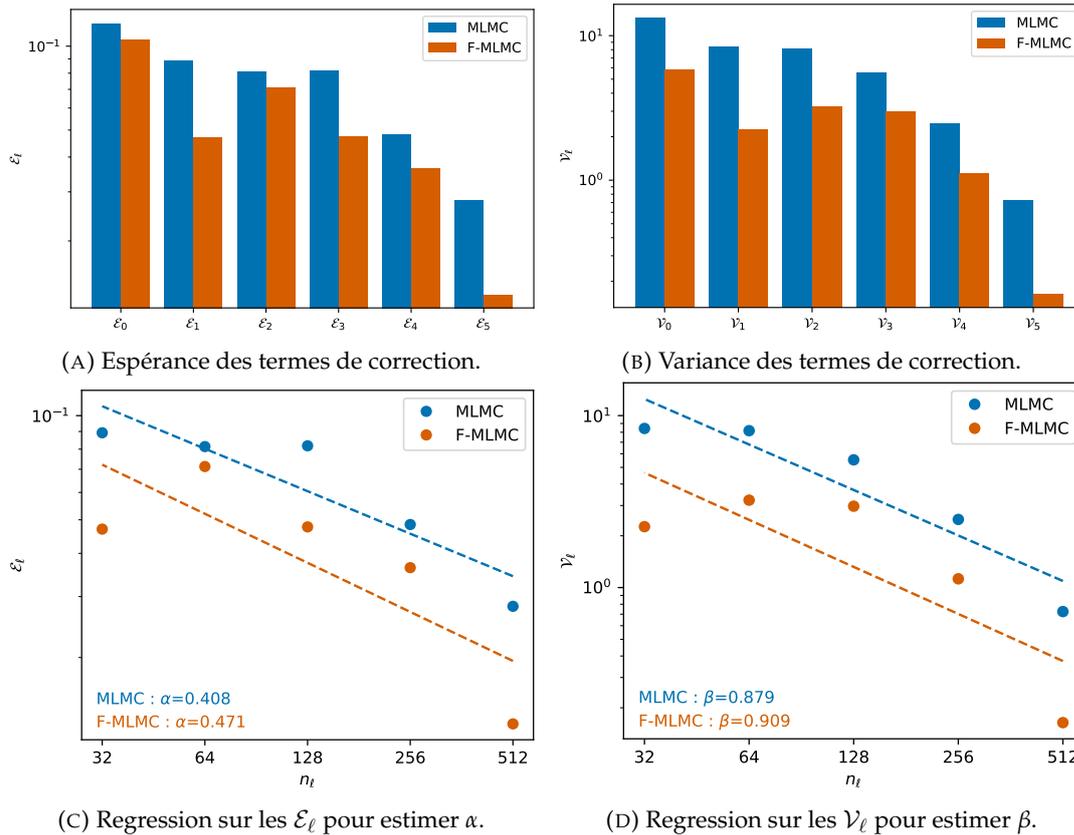


FIGURE 4.7 – Estimation des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  (cf. équation (3.109)) pour  $\ell = 0, \dots, L$  (haut) et regression sur ces quantités afin d'estimer  $\alpha$  et  $\beta$  (bas) pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006}\mathbf{X}_L]$ . Les estimations sont faites à partir de 1000 échantillons.

L'ajout des filtres ne permet pas, ici, de faire évoluer les espérances ou variances des termes de correction pour se rapprocher des hypothèses du théorème de convergence du MLMC. Les grilles les plus grossières semblent encore détériorer la qualité de l'estimateur et, bien que les quantités  $\alpha$  et  $\beta$  soient légèrement plus élevées pour les estimateurs F-MLMC, la décroissance clairement non linéaire des  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  ne permet pas de conclure quant à la MSE du F-MLMC comparé au MLMC. Il faut, cependant, noter que les  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  sont constamment plus faibles pour le F-MLMC que pour le MLMC. La détérioration de la variance des grilles les plus grossières seraient alors moins important pour le F-MLMC. C'est effectivement ce que l'on observe avec la variance totale des estimateurs sur la figure 4.8.

L'ajout des filtres permet de réduire la variance totale de l'estimateur dans tous les cas et, là où l'ajout des niveaux les plus grossières détériore fortement la variance du MLMC, leur impact sur le F-MLMC est faible. Les derniers niveaux n'apportent rien à l'estimateur mais les filtres semblent réduire une grande partie de leurs effets

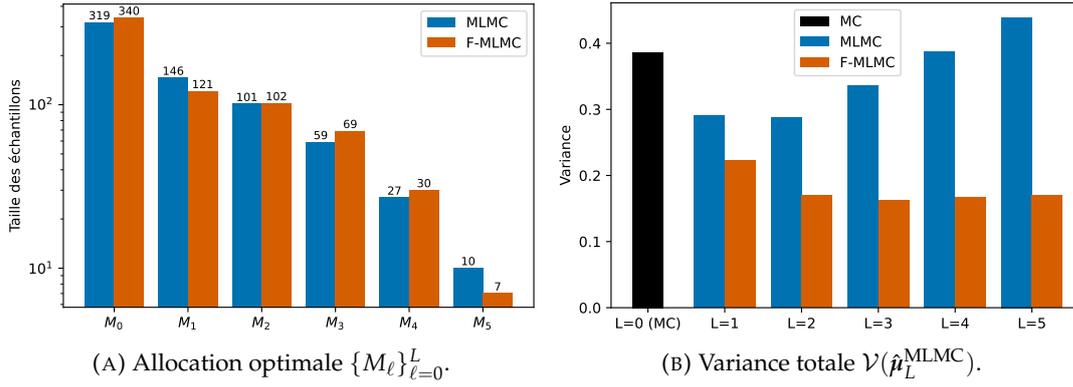
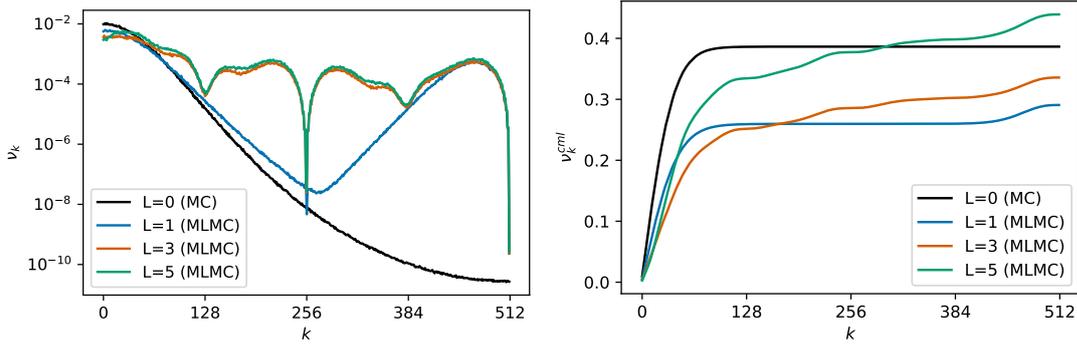


FIGURE 4.8 – Allocation optimale des échantillons pour les estimateurs MLMC et F-MLMC à 6 niveaux, et la variance totale résultante des estimateurs MC ( $L = 0$ ) et de différents estimateurs MLMC et F-MLMC pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

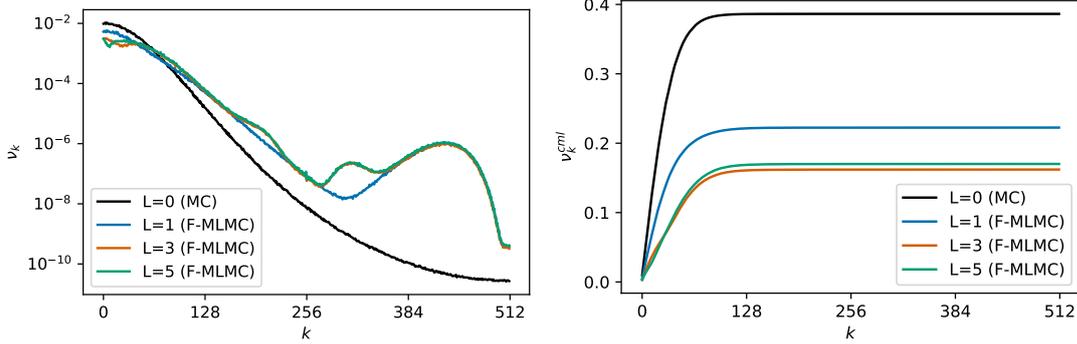
négatifs, même avec une allocation optimale similaire entre MLMC et F-MLMC. Les meilleurs estimateurs sont ceux à 3 niveaux ( $L = 2$ ) où les filtres apportent une réduction conséquente de la variance totale (environ 30% de réduction). En plus d'améliorer la variance de l'estimateur, les filtres semblent contribuer à l'amélioration de la stabilité de l'estimateur face à l'utilisation de niveaux non-adaptés au problème considéré. Finalement, la figure 4.9 présente la décomposition spectrale de la variance des estimateurs MLMC et F-MLMC pour  $L \in \{1, 3, 5\}$ .

La réduction de la variance sur l'ensemble du spectre pour le F-MLMC est ici bien visible. Pour les estimateurs MLMC basés sur 4 et 6 niveaux, on observe sur les courbes de variance cumulée que l'entièreté du spectre contribue, de manière significative, à l'augmentation de la variance totale. Pour le F-MLMC on retrouve des courbes similaires aux cas précédents où les composantes associées aux petites et moyennes échelles sont négligeables. La majorité de la variance totale provient des grandes échelles et cette variance a été également réduite par l'étape de filtrage. On remarque que le F-MLMC à 6 niveaux reste moins bon, en terme de variance totale, que celui à 4 niveaux. Cependant, la différence de variance est bien moindre comparée à la différence entre la variance du MLMC à 6 et 4 niveaux. Le filtrage des petites échelles qui détériorent la variance rend l'estimateur plus robuste à l'ajout de niveaux trop grossiers. Pour conclure cette illustration de l'estimateur F-MLMC, on présente la variance totale de l'estimateur F-MLMC\* basé sur la hiérarchie des modèles  $\tilde{f}_\ell^*$  de l'équation (4.41). La figure 4.10 compare la variance totale des estimateurs F-MLMC\* avec celle des estimateurs F-MLMC utilisés précédemment.

Les modèles  $\tilde{f}_\ell^*$  contiennent plus d'applications des opérateurs de filtrage  $\mathbf{S}_\ell$ , ce qui conduit à une réduction plus importante des petites échelles et de leur impact négatif sur la variance de l'estimateur. Cependant, on a vu que les filtres réduisaient également une partie des composantes associées aux grandes et moyennes échelles. Un filtrage trop important peut donc conduire à des pertes d'informations qui sont trop importantes comparées à la réduction des erreurs dues aux fines échelles. On observe ici que la variance totale des estimateurs F-MLMC\* est toujours légèrement plus élevée que celle des estimateurs F-MLMC, et ce pour les deux problèmes considérés. L'étape supplémentaire de filtrage dans la définition des opérateurs ne semble pas rentable ici.

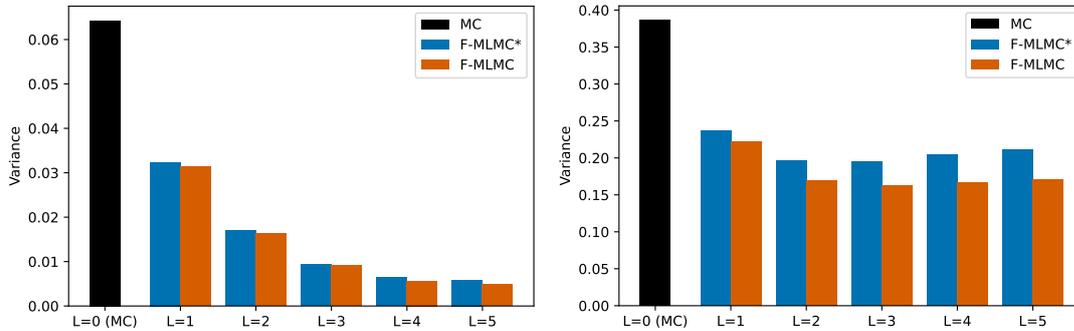


(A) Variance spectrale  $\nu$  et variance cumulée  $\nu^{\text{cml}}$  des estimateurs MLMC.



(B) Variance spectrale  $\nu$  et variance cumulée  $\nu^{\text{cml}}$  des estimateurs F-MLMC.

FIGURE 4.9 – Variance spectrale et cumulée de l’estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC (haut) et F-MLMC (bas), avec  $L \in \{1, 3, 5\}$ , pour le problème d’estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.



(A) Variance totale pour l’estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . (B) Variance totale pour l’estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ .

FIGURE 4.10 – Variance totale des estimateurs F-MLMC\* (bleu) et F-MLMC (orange) pour l’estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  (gauche) et de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  (droite). Les variances des estimateurs F-MLMC sont exactement les mêmes que précédemment et sont ici affichées pour faciliter la comparaison. Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.





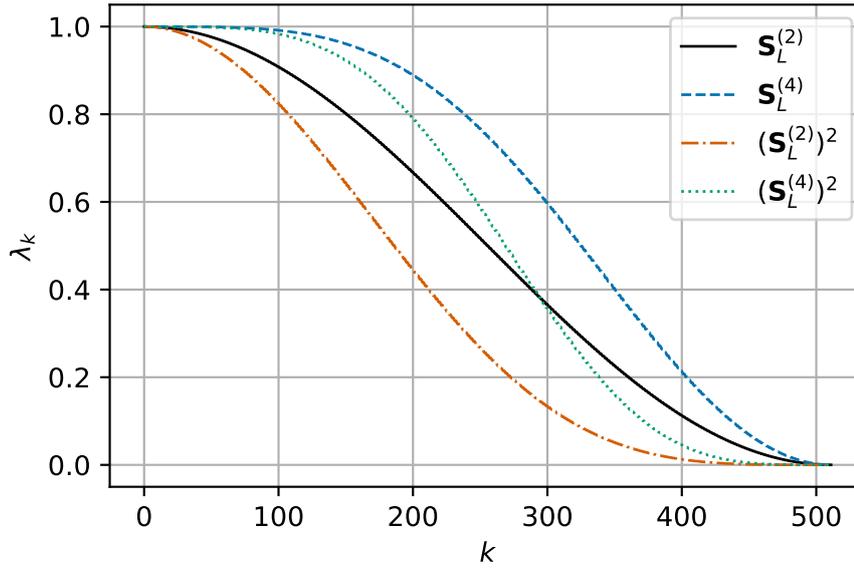
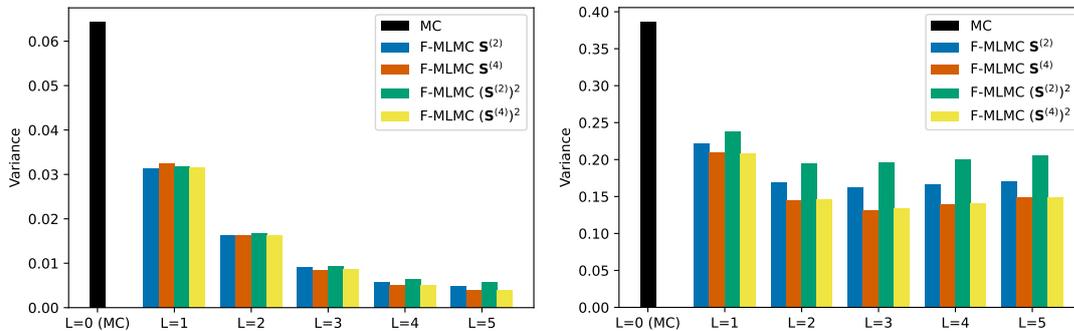


FIGURE 4.11 – Les valeurs propres des matrices  $\mathbf{S}_L^{(2)}$ ,  $\mathbf{S}_L^{(4)}$ ,  $(\mathbf{S}_L^{(2)})^2$  et  $(\mathbf{S}_L^{(4)})^2$ , toutes de taille  $512 \times 512$ , dans la base de Hartley réordonnée  $\mathbf{H}_L^*$ .

Finalement le filtre d'ordre 4 itéré possède une coupure plus nette entre la conservation des grandes échelles et le filtrage des fines échelles. Il est compliqué, à l'avance, de savoir lequel de ces filtres permettra la meilleure réduction de variance et on propose donc de les tester sur l'illustration 1D utilisée dans les sections 3.4 et 4.3 afin d'observer les différences sur l'estimateur MLMC. Les variances totales des différents estimateurs F-MLMC basés sur les différents filtres sont présentées dans la figure 4.12.

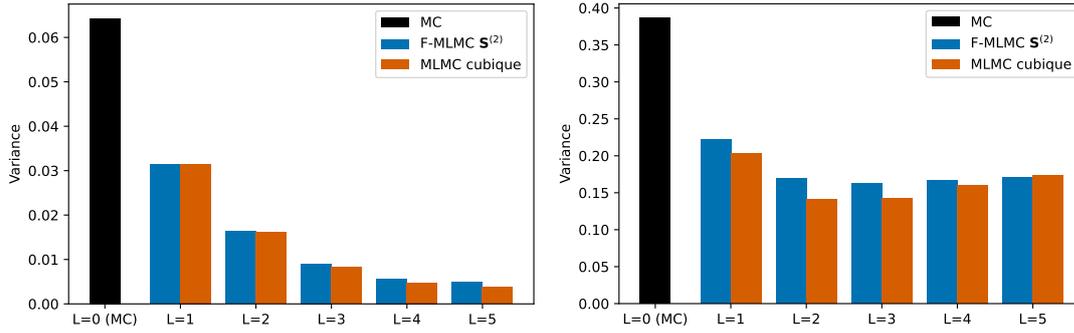


(A) Variance totale pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . (B) Variance totale pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ .

FIGURE 4.12 – Variance totale des estimateurs F-MLMC utilisant les filtres  $\mathbf{S}_\ell^{(2)}$  (bleu),  $\mathbf{S}_\ell^{(4)}$  (orange),  $(\mathbf{S}_\ell^{(2)})^2$  (vert) et  $(\mathbf{S}_\ell^{(4)})^2$  (jaune) pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  (gauche) et de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  (droite). Les variances des estimateurs F-MLMC basés sur le filtre  $\mathbf{S}_\ell^{(2)}$  sont exactement les mêmes que précédemment et sont ici affichées pour faciliter la comparaison. Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

Pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ , les résultats des différents estimateurs sont très similaires, le filtre de Shapiro d'ordre 4 (itéré ou non) semble apporter une légère réduction de la variance pour les estimateurs avec  $L \geq 3$ . Dans le cas de l'estimation





(A) Variance totale pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . (B) Variance totale pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ .

FIGURE 4.13 – Variance totale des estimateurs F-MLMC utilisant les filtres  $\mathbf{S}_\ell^{(2)}$  (bleu) et des estimateurs MLMC basés sur l'interpolation cubique (orange) pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  (gauche) et de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  (droite). Les variances des estimateurs F-MLMC basés sur le filtre  $\mathbf{S}_\ell^{(2)}$  sont les mêmes que précédemment et sont ici affichées pour faciliter la comparaison. Le niveau le plus fin  $L$  correspond à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

convergence [33, 17]. Ces quelques expériences confirment le besoin d'effectuer une étude pour exprimer plus clairement les différences entre les opérateurs comparés ici, ce qui pourrait permettre d'aboutir à une meilleure compréhension du lien entre ordre des opérateurs de transfert et variance de l'estimateur. En effet, tester différents filtres pour trouver celui apportant la meilleure réduction de variance peut être une étape longue et coûteuse, il serait donc intéressant d'avoir des méthodes permettant de guider ou d'aider au choix de ces opérateurs.

### 4.4.3 Weighted MLMC

#### Liens entre F-MLMC et WMLMC

Lors du choix des opérateurs de prolongation et de restriction des équations (3.19) et (3.20), plusieurs propriétés ont été imposées sur ces opérateurs dans les équations (3.15) et (3.17). La seconde propriété portait sur le fait que la restriction d'un bruit blanc devait rester un bruit blanc. Cette propriété, couplée à la contrainte  $\mathbf{R}_{\ell'}^\ell = \alpha^{\ell'-\ell} (\mathbf{P}_\ell^{\ell'})^\top$ , imposait de choisir  $\alpha = 2^{-1/2}$ . Cependant, dès que l'on ajoute une étape de pré-filtrage et que l'on considère l'opérateur de restriction  $\bar{\mathbf{R}}_\ell^{\ell-1}$  de l'équation (4.1), cette propriété n'est plus vérifiable pour  $\alpha \in \mathbb{R}$  car  $\bar{\mathbf{R}}_\ell^{\ell-1} (\bar{\mathbf{R}}_\ell^{\ell-1})^\top \neq \mathbf{I}_{n_{\ell-1}}$ . La propriété que l'on impose alors pour fixer le facteur  $\alpha$  est, pour tout  $\ell$  et  $\ell'$  tels que  $n_{\ell'} > n_\ell$ ,

$$\bar{\mathbf{R}}_{\ell'}^\ell (\bar{\mathbf{R}}_{\ell'}^\ell)^\top \mathbf{1}_{n_\ell} = \mathbf{1}_{n_\ell} \quad (4.46)$$

où  $\mathbf{1}_{n_\ell} \in \mathbb{R}^{n_\ell}$  est le vecteur constant égal à 1. Cette propriété peut s'interpréter comme la conservation des vecteurs constants par application de  $\bar{\mathbf{R}}_{\ell'}^\ell (\bar{\mathbf{R}}_{\ell'}^\ell)^\top$ . On remarque que cette propriété est bien vérifiée pour les opérateurs de restriction sans filtres de l'équation (3.20) utilisés par l'estimateur MLMC. Pour les opérateurs filtrés  $\bar{\mathbf{R}}_\ell^{\ell-1}$  où

$\ell = 1, \dots, L$ , on a

$$\bar{\mathbf{R}}_\ell^{\ell-1} (\bar{\mathbf{R}}_\ell^{\ell-1})^\top \mathbf{1}_{n_\ell} = \mathbf{R}_\ell^{\ell-1} \mathbf{S} \mathbf{S}^\top (\mathbf{R}_\ell^{\ell-1})^\top \mathbf{1}_{n_\ell} \quad (4.47)$$

$$= \frac{1}{32} \begin{bmatrix} 10 & 10 & 5 & 1 & & & 1 & 5 \\ 5 & 10 & 10 & 5 & 1 & & & 1 \\ 1 & 5 & 10 & 10 & 5 & 1 & & \\ & & & & \ddots & & & \\ & & & & & & & \end{bmatrix} \mathbf{1}_{n_\ell} = \mathbf{1}_{n_\ell} \quad (4.48)$$

Le choix du facteur  $\alpha = 2^{-1/2}$  est donc également celui qui permet de vérifier cette propriété pour les opérateurs  $\bar{\mathbf{R}}_\ell^\ell$  définis à partir de l'équation (4.1) comme combinaison d'un filtre de Shapiro d'ordre 2 (4.5) et de la restriction (3.20). On aimerait, cependant, vérifier si le choix d'imposer la propriété (4.46), et donc le choix du facteur  $\alpha = 2^{-1/2}$ , est bien celui qui minimise la variance de l'estimateur F-MLMC. Par linéarité des opérateurs utilisés et des modèles considérés, on a linéarité des opérateurs  $\bar{f}_\ell$ , pour  $\ell = 0, \dots, L$ , utilisés par le F-MLMC, et donc

$$\bar{f}_\ell = \alpha^{L-\ell} \bar{\mathbf{P}}_{\ell-1}^L \mathbf{A}_\ell (\bar{\mathbf{P}}_\ell^L)^\top. \quad (4.49)$$

En notant  $\hat{f}_\ell := \bar{\mathbf{P}}_{\ell-1}^L \mathbf{A}_\ell (\bar{\mathbf{P}}_\ell^L)^\top$  on peut réécrire l'estimateur F-MLMC de l'équation (4.4) comme

$$\hat{\boldsymbol{\mu}}_L^{\text{F-MLMC}} = \frac{1}{M_0} \sum_{i=1}^{M_0} \alpha^L \hat{f}_0(\mathbf{X}_L^{(0,i)}) + \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \alpha^{L-\ell} \hat{f}_\ell(\mathbf{X}_L^{(\ell,i)}) - \alpha^{L-\ell+1} \hat{f}_{\ell-1}(\mathbf{X}_L^{(\ell,i)}). \quad (4.50)$$

En choisissant  $\alpha = 2^{-1/2}$  on retrouve bien l'estimateur utilisé précédemment. Sous cette forme, l'estimateur F-MLMC apparaît similaire à un estimateur weighted MLMC (WMLMC) [56, 66]. On pose l'estimateur WMLMC de  $\mathbb{E}[f_L(\mathbf{X}_L)]$

$$\hat{\boldsymbol{\mu}}_L^{\text{WMLMC}} = \frac{1}{M_0} \sum_{i=1}^{M_0} \beta_0 \hat{f}_0(\mathbf{X}_L^{(0,i)}) + \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \beta_\ell \hat{f}_\ell(\mathbf{X}_L^{(\ell,i)}) - \beta_{\ell-1} \hat{f}_{\ell-1}(\mathbf{X}_L^{(\ell,i)}), \quad (4.51)$$

avec  $L+1$   $M_\ell$ -échantillons indépendants  $\{\mathbf{X}_L^{(i)}\}_{i=1}^{M_\ell}$  de  $\mathbf{X}_L$  et  $\{\beta_\ell\}_{\ell=0}^L$  une suite de poids réels. Sous la condition  $\beta_L = 1$ , on a bien la somme télescopique

$$\mathbb{E}[\hat{\boldsymbol{\mu}}_L^{\text{WMLMC}}] = \beta_0 \mathbb{E}[\hat{f}_0(\mathbf{X}_L)] + \sum_{\ell=1}^L \beta_\ell \mathbb{E}[\hat{f}_\ell(\mathbf{X}_L)] - \beta_{\ell-1} \mathbb{E}[\hat{f}_{\ell-1}(\mathbf{X}_L)] = \mathbb{E}[f_L(\mathbf{X}_L)], \quad (4.52)$$

et le biais de cet estimateur est bien le biais du plus haut niveau de fidélité  $L$ . On souhaite maintenant trouver les poids  $\{\beta_\ell\}_{\ell=0}^{L-1}$  qui minimisent  $\mathbb{V}[\hat{\boldsymbol{\mu}}_L^{\text{WMLMC}}]$ . On reprend les formulations des estimateurs MLBLUE [63, 62] pour déterminer la valeur optimale des poids  $\{\beta_\ell\}_{\ell=0}^{L-1}$ . En reprenant les résultats de [63, 18], on obtient l'expression suivante des poids qui minimisent la variance de l'estimateur

$$\beta_\ell = M_\ell \mathbf{C}_\ell^{-1} \mathbf{J}_\ell \left( \sum_{\ell'=0}^L M_{\ell'} \mathbf{J}_{\ell'}^\top \mathbf{C}_{\ell'}^{-1} \mathbf{J}_{\ell'} \right)^{-1} \mathbf{e}_L, \quad (4.53)$$

où  $\mathbf{C}_0 := \mathbb{V}[\hat{f}_0(\mathbf{X}_L)] \in \mathbb{R}$  et  $\mathbf{C}_\ell := \mathbb{V}[[\hat{f}_\ell(\mathbf{X}_L), \hat{f}_{\ell-1}(\mathbf{X}_L)]^\top] \in \mathbb{R}^{2 \times 2}$  sont les matrices de variance-covariance associées aux différents niveaux,  $\mathbf{J}_0 \in \mathbb{R}^{1 \times L+1}$  est l'opérateur de sélection du niveau 0 et  $\mathbf{J}_\ell \in \mathbb{R}^{2 \times L+1}$  l'opérateur de sélection du couple  $(\ell, \ell-1)$ ,  $\mathbf{e}_L := (0 \ 0 \ \dots \ 1)^\top \in \mathbb{R}^{L+1}$  est le vecteur  $L+1$  de la base canonique de  $\mathbb{R}^{L+1}$ . On

pose le vecteur  $M = (M_0 \ M_1 \ \dots \ M_L)^T \in \mathbb{R}^{L+1}$  et on note

$$\Phi(M) := \sum_{\ell=0}^L M_\ell \mathbf{J}_\ell^T \mathbf{C}_\ell^{-1} \mathbf{J}_\ell. \quad (4.54)$$

Les poids optimaux  $\{\beta_\ell\}_{\ell=0}^{L-1}$  dépendent donc de l'allocation  $\{M_\ell\}_{\ell=0}^{L-1}$ , et celle allocation optimale des échantillons sur les différents niveaux s'obtient par résolution du problème suivant

$$\min_{M \in \mathbb{R}^{L+1}} \mathbb{V}[\hat{\mu}_L^{\text{WMLMC}}(M)], \quad \text{tel que} \quad \sum_{\ell=0}^L M_\ell (\mathcal{C}_\ell + \mathcal{C}_{\ell-1}) \leq \eta, \quad (4.55)$$

où  $\mathcal{C}_\ell$  est le coût de calcul moyen d'une évaluation de  $\hat{f}_\ell$  et  $\eta$  est le budget total imposé. En injectant l'expression des poids (4.53) dans l'expression de la variance du WMLMC, on obtient

$$\mathbb{V}[\hat{\mu}_L^{\text{WMLMC}}(M)] = \mathbf{e}_L^T \Phi(M)^{-1} \mathbf{e}_L. \quad (4.56)$$

Pour une hiérarchie de modèles  $\{\hat{f}_\ell\}_{\ell=0}^L$  et un budget total  $\eta$  fixé, il est alors possible d'obtenir l'allocation optimale par résolution numérique du problème (4.55), puis les poids optimaux avec l'équation (4.53). Les matrices de variance-covariance apparaissant dans les expressions des poids (4.53) sont généralement inconnues et doivent être estimées. L'estimation de ces matrices doit être indépendante de l'estimation de  $\mathbb{E}[f_L(\mathbf{X}_L)]$  pour éviter d'introduire du biais. La résolution numérique de problème d'allocation optimale et l'estimation des matrices de variance-covariance implique des coûts de calcul supplémentaire par rapport à un simple estimateur MLMC et différentes stratégies de résolution sont discutées dans [63, 14].

L'estimateur F-MLMC de l'équation (4.50) a donc la même forme qu'un estimateur WMLMC avec la suite de poids  $\{\alpha^{L-\ell}\}_{\ell=0}^L$ . Obtenir les poids optimaux  $\{\beta_\ell\}_{\ell=0}^L$  de l'estimateur WMLMC (4.51) peut donc guider le choix du paramètre  $\alpha$  qui minimise la variance de l'estimateur F-MLMC. La comparaison des poids optimaux obtenus avec les poids  $\{\alpha^{L-\ell}\}_{\ell=0}^L$  où  $\alpha = 2^{-1/2}$  est notamment intéressante pour vérifier si la propriété de l'équation (4.46) que l'on a imposé permet bien d'aboutir à des opérateurs de transfert qui minimisent la variance de l'estimateur F-MLMC.

### Illustration 1D WMLMC

En reprenant l'illustration 1D utilisée plus haut pour comparer le MLMC et le F-MLMC, il est possible d'utiliser le WMLMC et, en déterminant les poids optimaux, déduire le coefficient  $\alpha$  optimal pour l'opérateur de restriction utilisé. On reprend donc le même contexte que la section 4.3, l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  et  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  pour  $\mathbf{X}_L \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{I}_{n_L})$  avec le modèle fin  $f_L$  défini sur une grille de taille 512. Le budget total des estimateurs reste  $\eta = 1000\mathcal{C}_L$ . Les matrices de covariance  $\{\mathbf{C}_\ell\}_{\ell=0}^L$  sont estimées à partir d'échantillons pilotes de taille 10 000 et le problème d'allocation optimale sous contrainte de budget total (4.55) est résolu numériquement à l'aide d'un algorithme des régions de confiance pour de l'optimisation sous contraintes décrit dans [8]. Les poids optimaux  $\{\beta_\ell\}_{\ell=0}^{L-1}$  sont finalement obtenus avec l'équation (4.53). On considère d'abord les modèles basés sur les opérateurs de Galerkin

$$\tilde{f}_\ell := \mathbf{P}_\ell^L \mathbf{R}_L^\ell \mathbf{A}_L \mathbf{P}_\ell^L (\mathbf{P}_\ell^L)^T, \quad \ell = 0, \dots, L-1, \quad (4.57)$$

qui sont utilisés pour définir un estimateur WMLMC à partir de l'équation (4.51). On considère également les modèles

$$\bar{f}_\ell := \bar{\mathbf{P}}_L^L \mathbf{R}_L^\ell \mathbf{A}_L \mathbf{P}_\ell^L (\bar{\mathbf{P}}_\ell^L)^T \quad (4.58)$$

qui sont utilisés pour définir un estimateur F-WMLMC, toujours à partir de l'équation (4.51). Ces deux estimateurs correspondent exactement aux estimateurs MLMC et F-MLMC utilisés dans la section 4.3 si les poids sont  $\beta_\ell = \alpha^{L-\ell} = \sqrt{2}^{\ell-1}$ . On affiche, sous forme de tableau, les poids optimaux obtenus pour l'estimateur WMLMC et F-WMLMC pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  (tableau 4.1) et  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  (tableau 4.2).

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$\sqrt{2}^{\ell-5}$
0	0.6962	0.4804	0.3216	0.2023	0.1121	0.1768
1	1	0.7008	0.4852	0.3264	0.2035	0.25
2		1	0.7015	0.4872	0.3269	0.3536
3			1	0.7025	0.4874	0.5
4				1	0.7026	0.7071
5					1	1

(A) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les estimateurs WMLMC.

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur F-WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$\sqrt{2}^{\ell-5}$
0	0.7043	0.5000	0.3538	0.2497	0.1727	0.1768
1	1	0.7060	0.5005	0.3558	0.2525	0.25
2		1	0.7067	0.5013	0.3565	0.3536
3			1	0.7072	0.5015	0.5
4				1	0.7073	0.7071
5					1	1

(B) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les estimateurs F-WMLMC.

TABLE 4.1 – Tables des poids optimaux  $\{\beta_\ell\}_{\ell=0}^L$  des estimateurs WMLMC et F-WMLMC pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$  et pour différentes valeurs de  $L$ . Les poids sont obtenus à partir de l'équation (4.53) après résolution du problème (4.55) et avec estimation des covariances à partir d'un échantillon pilote de taille 10 000.

Dans ces tableaux, chaque colonne correspond aux poids obtenus pour un estimateur avec un nombre de niveaux  $L$  fixé. Le poids du niveau le plus fin  $\ell = L$  est toujours fixé à 1 pour assurer la somme télescopique (4.52). La comparaison avec les poids  $\sqrt{2}^{\ell-5}$  utilisés par le F-MLMC de la section 4.3 est à faire en considérant les diagonales du tableau, c'est-à-dire que l'on compare, par exemple, le poids  $\sqrt{2}^{4-5} = 1/\sqrt{2}$  avec le poids numéroté par  $\ell = L - 1$  de chaque estimateur. Pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ , les poids obtenus et présentés dans le tableau 4.1 sont très proches des facteurs  $\sqrt{2}^{\ell-L}$ . Cela confirme que l'utilisation d'un facteur  $1/\sqrt{2}$  devant l'opérateur de restriction est le bon choix, que ce soit pour l'opérateur sans ou avec le filtre de Shapiro d'ordre 2. On observe cependant des poids légèrement plus faibles que ceux attendus pour le dernier niveau de l'estimateur WMLMC avec  $L = 4$  et  $L = 5$ , ce qui semble indiquer que ces niveaux sont moins

utiles. Afin de vérifier si les poids obtenus sont bien ceux qui minimisent la variance de l'estimateur, on affiche les variances totales de ces estimateurs dans la figure 4.14 où elles sont comparées aux variances des estimateurs MLMC et F-MLMC obtenues précédemment.

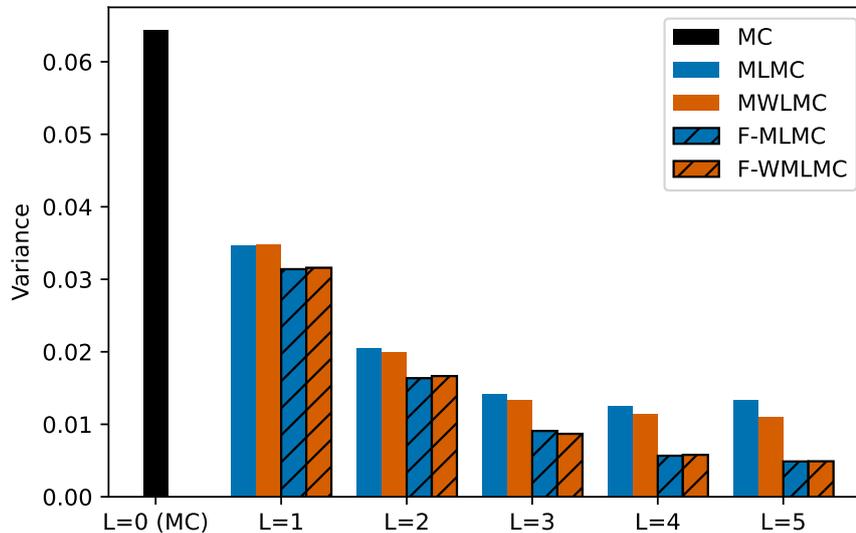


FIGURE 4.14 – Variance totale des estimateurs MLMC, WMLMC, F-MLMC et F-WMLMC pour  $L \in \{1, \dots, 5\}$  ainsi que l'estimateur MC ( $L = 0$ ) pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{006} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

On constate, premièrement, que les profils de variance sont très proches entre les estimateurs MLMC et WMLMC, ainsi qu'entre les estimateurs F-MLMC et F-WMLMC. En théorie, l'ajout des poids optimaux ne peut pas, par définition, conduire à une augmentation de la variance totale de l'estimateur. En pratique, les différentes estimations et erreurs numériques peuvent conduire à des résultats légèrement différents. Ici, la similarité des variances totales, ainsi que les poids optimaux obtenus qui conduisent à un estimateur WMLMC quasiment équivalent au MLMC semble indiquer que l'estimateur MLMC utilisé est déjà de variance minimale pour le choix du coefficient  $\alpha$ . Le même raisonnement peut être émis pour les estimateurs F-MLMC et F-WMLMC. Pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ , les poids optimaux obtenus sont reportés dans le tableau 4.2

Contrairement aux résultats précédents, les poids optimaux de l'estimateur WMLMC pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  sont significativement différents des poids utilisés pour l'estimateur MLMC. Le poids du dernier niveau de l'estimateur  $L = 5$ , par exemple, est presque 10 fois plus petit que celui utilisé par le MLMC. On rappelle, cependant, que pour ce problème, la figure 4.8 avait montré que l'ajout de niveaux passé le troisième détériorait la variance totale de l'estimateur. Il est donc cohérent, avec le WMLMC, d'obtenir des poids plus faibles que ceux utilisés précédemment, notamment pour les niveaux les plus grossiers. Les poids du F-WMLMC sont, eux, bien plus similaires à ceux utilisés pour le F-MLMC, même si l'on observe encore quelques différences significatives pour les poids associées aux derniers niveaux. La figure 4.15 présente les variances totales de ces différents estimateurs.

Les estimateurs F-MLMC et F-WMLMC ont des variances totales très similaires, ce qui est cohérent avec les poids optimaux obtenus qui sont, eux aussi, très proches.

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$\sqrt{2}^{\ell-5}$
0	0.6398	0.3728	0.1856	0.0714	0.0238	0.1768
1	1	0.6467	0.3680	0.1624	0.0520	0.25
2		1	0.6439	0.3553	0.1524	0.3536
3			1	0.6403	0.3510	0.5
4				1	0.6388	0.7071
5					1	1

(A) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les estimateurs WMLMC.

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur F-WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$\sqrt{2}^{\ell-5}$
0	0.7043	0.4940	0.3259	0.2017	0.1312	0.1768
1	1	0.7116	0.4930	0.3210	0.2014	0.25
2		1	0.7096	0.4911	0.3230	0.3536
3			1	0.7091	0.4945	0.5
4				1	0.7117	0.7071
5					1	1

(B) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les estimateurs F-WMLMC.

TABLE 4.2 – Tables des poids optimaux  $\{\beta_\ell\}_{\ell=0}^L$  des estimateurs WMLMC et F-WMLMC pour le problème d'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  et pour différentes valeurs de  $L$ . Les poids sont obtenus à partir de l'équation (4.53) après résolution du problème (4.55) et avec estimation des covariances à partir d'un échantillon pilote de taille 10 000.

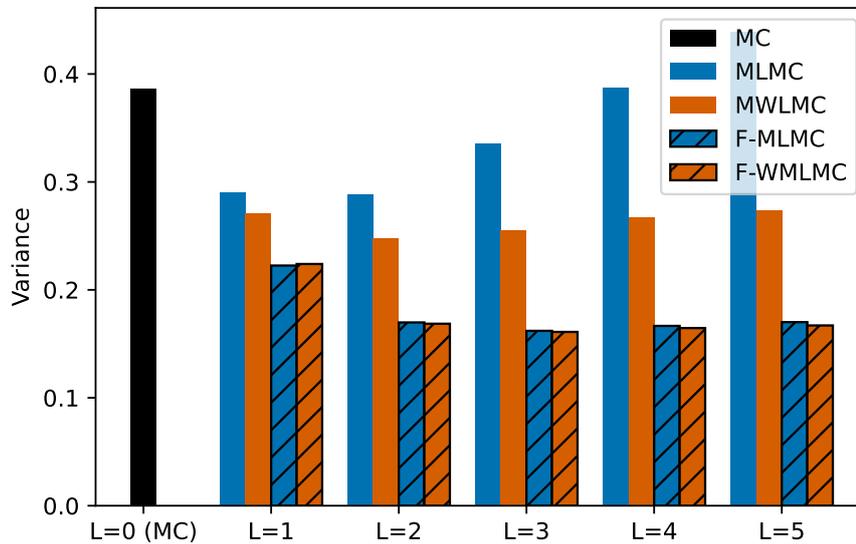


FIGURE 4.15 – Variance totale des estimateurs MLMC, WMLMC, F-MLMC et F-WMLMC pour  $L \in \{1, \dots, 5\}$  ainsi que l'estimateur MC ( $L = 0$ ) pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

Pour la version sans filtres, il est intéressant de constater que l'estimateur WMLMC est bien meilleur que le MLMC. Les poids obtenus semblent compenser l'ajout des

niveaux les plus grossiers pour lesquels la représentation du champ considéré est complètement polluée par les effets négatifs des opérateurs de transfert.

Le facteur  $2^{-1/2}$  ajouté devant les opérateurs de restriction semble être optimal dans les cas considérés ici, excepté le cas du MLMC pour l'estimation de  $\mathbb{E}[\mathbf{A}_L^{001} \mathbf{X}_L]$  où les niveaux les plus grossiers détériorent la variance totale. La concordance des poids obtenus entre le F-MLMC et le F-WMLMC est notamment intéressante car, dans le cas des opérateurs filtrés  $\bar{\mathbf{R}}_\ell^{\ell-1}$ , les expériences faites permettent de conforter notre choix d'imposer la propriété (4.46) sur les opérateurs de transfert. Le choix de ce facteur  $2^{-1/2}$  est justifié et les illustrations faites jusqu'à maintenant correspondent au cas optimal. L'utilisation du WMLMC peut donc aider au choix des opérateurs de transfert en permettant de trouver le facteur optimal à appliquer devant ces opérateurs.

## 4.5 Conclusion

Ce chapitre a introduit l'estimateur F-MLMC (4.4) qui ajoute une étape de filtrage avant la restriction et après la prolongation d'un champ discrétisé. L'objectif est d'atténuer les effets négatifs des opérateurs de transferts sur l'estimateur en filtrant les fines échelles qui sont mal représentées sur les grilles grossières et polluent la représentation du champ. Dans le cas simplifié de modèles linéaires, symétriques et circulants, l'expression de la variance des estimateurs MLMC et F-MLMC à deux niveaux permet une analyse spectrale de l'effet des filtres. Le filtrage des fines échelles avec un filtre de Shapiro d'ordre 2 permet de réduire fortement les termes parasites introduits par les opérateurs de transfert qui apparaissent dans la variance de l'estimateur, ce qui se fait au prix d'une perte d'information sur les moyennes et grandes échelles. Les comparaisons de l'estimateur MLMC et F-MLMC sur l'illustration 1D étudiée montre une réduction significative de la variance de toutes les composantes spectrales, et donc une réduction significative de la variance totale de l'estimateur. La question du choix des filtres et des opérateurs de transfert qui minimisent la variance du F-MLMC se pose alors. Les expériences effectuées semblent montrer que, de manière similaire aux méthodes multigrilles, les opérateurs de transfert utilisés dans l'estimateur MLMC doivent être d'ordre suffisamment élevé et qu'il est possible d'augmenter l'ordre de ces opérateurs de transfert en ajoutant des filtres de Shapiro comme cela a été fait dans [17]. Une étude plus poussée sur l'ordre des opérateurs de transfert et leur lien avec la variance de l'estimateur pourrait confirmer un tel résultat. Outre tester simplement des opérateurs de filtrage d'ordre différents, il est également possible d'utiliser des méthodes de type MBLBLUE pour guider le choix des filtres. L'utilisation du WMLMC a ici permis, par exemple, de confirmer que l'utilisation du facteur  $2^{-1/2}$  devant les opérateurs de restriction était justifiée et même optimale dans certaines expériences. En utilisant l'extension du MBLBLUE présentée dans [18, section 4.4] pour des champs discrétisés incluant un changement de base, il peut être possible de trouver les poids optimaux à appliquer à chaque composante spectrale d'un champ pour minimiser la variance de l'estimateur. Obtenir des poids à appliquer à chaque composante spectrale reviendrait à trouver un filtre à appliquer aux champs obtenus après l'étape de prolongation, donc trouver un opérateur de post-filtrage qui minimise la variance. Une telle étude pourrait apporter une meilleure compréhension des filtres recherchés pour minimiser la variance de l'estimateur MLMC. Il est, cependant, important de rappeler que les développements théoriques effectués et l'illustration étudiée dans ce chapitre sont des cas extrêmement simplifiés qui ne représentent pas la complexité d'une application réelle. En effet, on a considéré des

simulateurs linéaires, symétriques et circulants, et l'espérance du champ discrétisé que l'on souhaite estimer est connue et égale au champ nul. L'objectif du chapitre suivant est donc de proposer un cas d'application plus complexe afin d'illustrer les améliorations apportées par l'estimateur F-MLMC.

## Chapitre 5

# Expériences numériques

Les chapitres précédents ont permis une meilleure compréhension de l'impact des opérateurs de transfert sur un estimateur multi-niveaux et ont démontré les avantages apportés par l'ajout des filtres sur un problème d'estimation de l'espérance, constante et nulle, d'un champ aléatoire discrétisé issu de l'application d'un modèle linéaire, symétrique et circulant à un champ aléatoire discrétisé d'entrée. Un de nos objectifs est d'étendre les applications du MLMC aux géosciences, et l'illustration de la section 4.3 n'est pas vraiment représentative des problèmes rencontrés dans ce domaine. Ce chapitre se consacre donc à l'étude d'un problème plus complexe basé sur des simulateurs non-linéaires, en 1D et en 2D. Les expériences numériques présentées ici permettent de comparer des performances des estimateurs MLMC et F-MLMC à celle d'un simple estimateur MC afin de vérifier, expérimentalement, si les conclusions des chapitres précédents se confirment.

### 5.1 Contexte et présentation du problème

#### 5.1.1 Problème de normalisation d'un opérateur de covariance basé sur la diffusion

La modélisation et l'estimation de matrices de covariance sont des étapes cruciales pour l'assimilation de données en géosciences [10], notamment pour la matrice de covariance de l'erreur d'ébauche  $\mathbf{B} \in \mathbb{R}^{n \times n}$ . Une approche générale pour modéliser de telles matrices de covariance est d'utiliser un opérateur linéaire différentiel discrétisé. On pense, par exemple, aux opérateurs de diffusion discrétisés qui permettent de représenter des matrices de noyau de covariance Matérn [75, 29]. L'utilisation de ces opérateurs de diffusion pour modéliser des covariances se retrouve notamment en assimilation de données pour l'océan [70], l'atmosphère [46] ou en géostatistique [58]. De tels opérateurs permettent de modéliser une structure de covariance, mais une étape de normalisation est ensuite nécessaire afin de pouvoir imposer les variances souhaitées. Autrement dit, si l'on note  $\mathbf{L} \in \mathbb{R}^{n \times n}$  une matrice de covariance modélisée comme un opérateur de diffusion, l'étape de normalisation consiste à construire la matrice de corrélation

$$\mathbf{C} := \mathbf{L}\mathbf{L}\mathbf{L}, \quad (5.1)$$

où

$$\mathbf{\Gamma} := \text{Diag}(\gamma_1, \dots, \gamma_n) = \text{Diag}(\text{diag}(\mathbf{L}))^{-1/2}. \quad (5.2)$$

Dans cette expression, l'opérateur  $\text{Diag}(\cdot)$  transforme un vecteur en une matrice diagonale dont les éléments diagonaux sont les entrées du vecteur, et l'opérateur  $\text{diag}(\cdot)$  retourne le vecteur contenant la diagonale de la matrice d'entrée. Une fois cette étape effectuée, il est facile de construire une matrice de covariance  $\mathbf{B}$  avec  $\mathbf{\Sigma}^2 = \mathbf{\Sigma}\mathbf{\Sigma} = \text{Diag}(\sigma_1^2, \dots, \sigma_n^2)$  la matrice diagonale des variances que l'on souhaite

imposer,

$$\mathbf{B} = \mathbf{\Sigma}\mathbf{C}\mathbf{\Sigma} = \mathbf{\Sigma}\mathbf{\Gamma}\mathbf{L}\mathbf{\Gamma}\mathbf{\Sigma}. \quad (5.3)$$

En pratique, quand la dimension  $n$  du problème est grande ( $n \sim 10^8$  dans des modèles opérationnels), l'opérateur de diffusion  $\mathbf{L}$ , qui est dense, n'est pas stocké explicitement. Seule sa représentation sous la forme d'une application à un vecteur est accessible. La diagonale de  $\mathbf{L}$ , qui sert à calculer la matrice  $\mathbf{\Gamma}$  des coefficients de normalisation, n'est pas directement accessible. Une méthode exacte pour obtenir cette diagonale est d'appliquer  $\mathbf{L}$  à tous les vecteurs  $\mathbf{e}_i$  de la base canonique de  $\mathbb{R}^n$ ,

$$\mathbf{v} := \text{diag}(\mathbf{L}) = ((\mathbf{L}\mathbf{e}_1)_1, (\mathbf{L}\mathbf{e}_2)_2, \dots, (\mathbf{L}\mathbf{e}_n)_n)^\top. \quad (5.4)$$

Cependant, cette méthode est prohibitivement coûteuse pour  $n$  grand. La méthode la plus utilisée actuellement [73] est d'estimer les coefficients de normalisation par randomisation avec un estimateur MC. Nous cherchons donc à appliquer l'estimateur MLMC à ce problème pour essayer d'améliorer l'estimation des coefficients de normalisation par rapport à l'estimateur MC. Avant de détailler l'estimateur MC utilisé, nous détaillons la construction de l'opérateur de diffusion  $\mathbf{L}$ .

### 5.1.2 Construction de l'opérateur de diffusion

Soit la fonction

$$\begin{aligned} u: \mathbb{R}^+ \times \mathcal{D} &\rightarrow \mathbb{R} \\ (t, x) &\mapsto u(t, x) \end{aligned} \quad (5.5)$$

telle que  $u(t, \cdot)$  soit de carré intégrable sur  $\mathcal{D} \subset \mathbb{R}^d$  avec  $d \in \{1, 2, 3\}$ . L'équation de diffusion s'écrit

$$\frac{\partial u}{\partial t} = \nabla \cdot (\kappa \nabla u), \quad (5.6)$$

avec  $t$  un pseudo-temps et  $\kappa: \mathcal{D} \rightarrow \mathbb{R}^{d \times d}$  un champ tensoriel, symétrique semi-défini positif, constant dans le temps, représentant la diffusivité. Cette équation est soumise à des conditions aux bords dépendantes de l'application considérée. Conformément à [50], on utilise un schéma d'Euler implicite en temps pour la résolution de cette équation. En discrétisant le temps avec un pas constant  $\Delta t$  et en notant  $t_k = k\Delta t$  pour  $k \in \mathbb{N}$ , on peut écrire l'approximation suivante de l'équation (5.6) :

$$\frac{u(t_k, x) - u(t_{k-1}, x)}{\Delta t} = \nabla \cdot (\kappa \nabla u(t_k, x)). \quad (5.7)$$

On s'intéresse à la solution de l'équation à un temps  $t_m$  avec  $m \in \mathbb{N}^*$ . On a donc

$$u(t_m, x) = (id - \Delta t \nabla \cdot \kappa \nabla)^{-m} u(t_0, x), \quad (5.8)$$

avec  $id$  la fonction identité. L'équation (5.8) représente donc un opérateur de diffusion implicite sur  $m$  pas de temps de longueur  $\Delta t$  appliqué à la condition initiale  $u(t_0, x)$ . La solution intégrale de cette équation sur  $\mathbb{R}^d$  définit un opérateur de covariance dont le noyau est une fonction Matérn [75, 29]. La méthode de résolution de cette équation est présentée dans [72, Section 3.2]. Dans la suite, on fixe un pas de temps unitaire  $\Delta t = 1$  et on impose  $m = 2q$  où  $q \in \mathbb{N}$ . De plus, on se restreint à un tenseur de diffusivité  $\kappa$  diagonal, i.e., à chaque point de l'espace est associée une matrice

$[K_{ij}]_{i,j=1,\dots,d}$  diagonale. Les termes diagonaux sont définis selon [71] suivant la relation

$$K_{ii}(x) = \frac{1}{2m - d - 2} D_{ii}^2(x), \quad (5.9)$$

avec  $m > (d + 2)/2$  et où  $D_{ii}(x)$  représente une longueur de corrélation dans la direction  $i$  au point  $x$  de l'espace pour  $i = 1, \dots, d$ . Le champ tensoriel qui associe à chaque point de l'espace la matrice  $[D_{i,j}]_{i,j=1,\dots,d}$ , ici considérée diagonale, est appelé champ de Daley [71].

On considère une discrétisation du domaine  $\mathcal{D}$  en une grille composée de  $n$  cellules dont  $\mathbf{W} \in \mathbb{R}^{n \times n}$  est la matrice de Gram, symétrique définie positive, qui encode sa structure. Les champs scalaires définis sur  $\mathcal{D}$  sont alors représentés par un vecteur de taille  $n$  dont les entrées correspondent à la valeur du champ au centre des cellules de la grille. On définit la matrice  $\Delta \in \mathbb{R}^{n \times n}$  qui représente la discrétisation de l'opérateur différentiel  $\nabla \cdot \kappa \nabla$  sur cette grille à l'aide d'une formulation forte. On impose que la discrétisation choisie soit telle que  $\Delta$  soit auto-adjoint par rapport au produit scalaire pondéré par  $\mathbf{W}$ , i.e.,  $\mathbf{W}\Delta = (\mathbf{W}\Delta)^T$ . On définit

$$\mathbf{L} := (\mathbf{I} - \Delta)^{-m} \mathbf{W}^{-1} \quad (5.10)$$

l'opérateur de diffusion associé à la discrétisation spatiale de l'équation (5.8). Cette matrice  $\mathbf{L}$  est symétrique car

$$(\Delta^{-m} \mathbf{W}^{-1})^T = \mathbf{W}^{-T} (\Delta^{-m})^T \quad (5.11)$$

$$= \mathbf{W}^{-T} \Delta^{-T} (\Delta^{-T})^{m-1} \quad (5.12)$$

$$= (\mathbf{W}\Delta)^{-T} (\Delta^{-T})^{m-1} \quad (5.13)$$

$$= (\mathbf{W}\Delta)^{-1} (\Delta^{-T})^{m-1} \quad (5.14)$$

$$= \Delta^{-1} \mathbf{W}^{-1} (\Delta^{-T})^{m-1} \quad (5.15)$$

$$= \dots \quad (5.16)$$

$$= \Delta^{-m} \mathbf{W}^{-1}. \quad (5.17)$$

Elle est également définie positive car  $-\Delta$  est définie positive.  $\mathbf{L}$  représente donc bien une matrice de covariance et admet une décomposition  $\mathbf{L} = \mathbf{A}\mathbf{A}^T$  avec  $\mathbf{A} \in \mathbb{R}^{n \times n}$ . Avec la décomposition supplémentaire  $\mathbf{W} = \mathbf{V}\mathbf{V}^T$  et en rappelant que  $m = 2q$ , on a

$$\mathbf{A} = (\mathbf{I} - \Delta)^{-q} (\mathbf{V}^{-1})^T. \quad (5.18)$$

Dans la suite, on désigne  $\mathbf{A}$  comme l'opérateur racine carrée de la diffusion. L'accès à cet opérateur  $\mathbf{A}$  peut également servir, après normalisation, à générer des champs gaussiens de variance unitaire et de corrélation  $\mathbf{C}$ . En effet, en partant de l'équation (5.1) on a

$$\mathbf{C} = \Gamma \mathbf{L} \Gamma = \Gamma \mathbf{A} \mathbf{A}^T \Gamma = (\Gamma \mathbf{A}) (\Gamma \mathbf{A})^T, \quad (5.19)$$

et donc, en prenant  $\varepsilon \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$ ,

$$\mathbf{Y} := \Gamma \mathbf{A} \varepsilon \sim \mathcal{N}(\mathbf{0}_n, \mathbf{C}). \quad (5.20)$$

En partant plutôt de l'équation (5.3), on peut générer des champs gaussiens de covariance  $\mathbf{B}$  où les variances peuvent être fixées librement avec la matrice diagonale  $\Sigma$ .

### 5.1.3 Estimation des coefficients de normalisation par randomisation

On revient maintenant sur l'estimation des coefficients de normalisation de l'opérateur de diffusion  $\mathbf{L}$  de l'équation (5.10). Plus spécifiquement, on veut estimer le vecteur  $\mathbf{v}$  des variances de  $\mathbf{L}$ , défini dans (5.4), qui servira par la suite à obtenir les coefficients de normalisation  $\gamma_i = v_i^{-1/2}$ . Soit  $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}_n, \mathbf{I}_n)$  un vecteur aléatoire de taille  $n$  suivant une distribution normale multivariée de moyenne nulle  $\mathbb{E}[\boldsymbol{\varepsilon}] = \mathbf{0}_n \in \mathbb{R}^n$  et de matrice de covariance identité  $\mathbb{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top] = \mathbf{I}_n \in \mathbb{R}^{n \times n}$ . On a alors les relations

$$\mathbf{L} = \mathbf{A}\mathbf{A}^\top = \mathbf{A}\mathbf{I}_n\mathbf{A}^\top = \mathbf{A}\mathbb{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top]\mathbf{A}^\top = \mathbb{E}[(\mathbf{A}\boldsymbol{\varepsilon})(\mathbf{A}\boldsymbol{\varepsilon})^\top] = \mathbf{C}[\mathbf{A}\boldsymbol{\varepsilon}], \quad (5.21)$$

et, par conséquent,

$$\mathbf{v} = \text{diag}(\mathbf{L}) = \mathbb{V}[\mathbf{A}\boldsymbol{\varepsilon}] = \mathbb{E}[(\mathbf{A}\boldsymbol{\varepsilon}) \odot (\mathbf{A}\boldsymbol{\varepsilon})], \quad (5.22)$$

où  $\mathbb{E}$  et  $\mathbb{V}$  s'appliquent élément par élément et  $\odot$  désigne le produit de Schur (élément par élément). Soit  $\{\boldsymbol{\varepsilon}^{(i)}\}_{i=1}^M$  un  $M$ -échantillon de  $\boldsymbol{\varepsilon}$ , on définit alors l'estimateur MC  $\hat{\mathbf{v}}$  de  $\mathbf{v}$  par

$$\hat{\mathbf{v}} = \frac{1}{M} \sum_{i=1}^M (\mathbf{A}\boldsymbol{\varepsilon}^{(i)}) \odot (\mathbf{A}\boldsymbol{\varepsilon}^{(i)}). \quad (5.23)$$

Finalement, on a

$$\hat{\mathbf{\Gamma}} := \text{diag}(\hat{v}_1^{-1/2}, \dots, \hat{v}_n^{-1/2}). \quad (5.24)$$

Cet estimateur permet d'obtenir une approximation des coefficients de normalisation avec seulement  $M$  applications de l'opérateur racine carrée de la diffusion,  $\mathbf{A}$ . De plus, un tel estimateur garantit que les variances obtenues sont positives. Cependant, la racine de l'erreur quadratique moyenne de l'estimateur MC est proportionnelle à  $M^{-1/2}$ . Une estimation précise requiert un grand nombre  $M$  d'applications et donc un coût de calcul important. Les estimateurs MLMC et F-MLMC présentés dans ce manuscrit sont appliqués à ce problème d'estimation des coefficients de normalisation d'un opérateur de covariance basé sur la diffusion et sont notamment comparés à la méthode d'estimation de ces coefficients par randomisation MC basée sur l'équation (5.23).

## 5.2 Normalisation 1D

On commence par considérer le problème d'estimation du vecteur  $\mathbf{v}$  de l'équation (5.4) dans un contexte 1D. On reprend un contexte similaire aux illustrations 1D utilisées dans les sections 3.4 et 4.3. Le domaine  $\mathcal{D} = [0, 1]$  est considéré périodique et est discrétisé par  $L + 1$  grilles régulières et centrées sur les cellules. Chaque grille est composée de  $n_\ell$  cellules avec la relation imposée  $n_\ell = 2^{\ell-L}n_L$  pour  $\ell = 0, \dots, L$  et pour un  $n_L$  fixé. Les matrices de Gram associées aux différentes grilles sont donc  $\mathbf{W}_\ell = n_\ell^{-1}\mathbf{I}_{n_\ell}$ .

On détaille ici la construction de la matrice  $\boldsymbol{\Delta} \in \mathbb{R}^{n_L \times n_L}$  qui correspond à la discrétisation de l'opérateur différentiel  $\nabla \cdot \boldsymbol{\kappa} \nabla$  de l'équation (5.8) sur la grille de taille  $n_L$  à l'aide des différences finies. On note, pour  $i = 0, \dots, n_L - 1$ ,  $x_i$  les points de l'espace qui correspondent aux centres des cellules de la grille et  $x_{i-\frac{1}{2}}$  les points de l'espace qui correspondent aux nœuds de la grille. Par périodicité du domaine, on a  $x_{n_L-\frac{1}{2}} = x_{-\frac{1}{2}}$ . On considère le vecteur  $\mathbf{u} = \{u(x_i)\}_{i=0}^{n_L-1}$  qui représente un champ discrétisé au centre des cellules de cette grille. Le tenseur de diffusivité  $\boldsymbol{\kappa}$  (cf. équation (5.9)) qui, à chaque point  $x$  de l'espace, associe une diffusivité  $K_{11}(x)$ , est

discrétisé aux interfaces entre les cellules de la grille, c'est-à-dire sur les nœuds de la grille. Pour  $i = 0, \dots, n_L - 1$ , on note  $K_{i-\frac{1}{2}} = K_{11}(x_{i-\frac{1}{2}})$ . Soit  $z(x) := \nabla \cdot \kappa \nabla u(x)$  et  $\mathbf{z} = \{z(x_i)\}_{i=0}^{n_L-1}$  le champ discrétisé correspondant. On cherche à obtenir une approximation  $\hat{\mathbf{z}}$  du vecteur  $\mathbf{z}$ . On utilise un schéma de différences finies centrées d'ordre 2 pour discrétiser  $\nabla \cdot \kappa \nabla$ , ce qui donne, pour  $i = 0, \dots, n_L - 1$

$$\hat{z}_i = (\Delta x)^{-2} \left( K_{i+\frac{1}{2}}[u_{i+1} - u_i] - K_{i-\frac{1}{2}}[u_i - u_{i-1}] \right) \quad (5.25)$$

$$= n_L^2 \left( u_{i-1}K_{i-\frac{1}{2}} - u_i[K_{i-\frac{1}{2}} + K_{i+\frac{1}{2}}] + u_{i+1}K_{i+\frac{1}{2}} \right), \quad (5.26)$$

avec  $\Delta x = n_L^{-1}$  le pas de discrétisation en espace. On définit alors la matrice

$$\Delta = n_L^2 \begin{bmatrix} -[K_{-\frac{1}{2}} + K_{\frac{1}{2}}] & K_{\frac{1}{2}} & & & K_{-\frac{1}{2}} \\ K_{\frac{1}{2}} & -[K_{\frac{1}{2}} + K_{\frac{3}{2}}] & K_{\frac{3}{2}} & & \\ & & \ddots & \ddots & \ddots \\ & & & K_{n_L-\frac{3}{2}} & -[K_{n_L-\frac{3}{2}} + K_{-\frac{1}{2}}] \\ K_{-\frac{1}{2}} & & & & \end{bmatrix} \quad (5.27)$$

qui est telle que  $\hat{\mathbf{z}} = \Delta \mathbf{u}$ . Dans cette illustration 1D on choisit de prendre un champ tensoriel de diffusion  $\kappa$  (5.9) homogène en espace, c'est-à-dire que l'on a  $D_{11}(x) = D \in \mathbb{R}$ , et donc  $K_{11}(x) = K \in \mathbb{R}$  pour tout  $x \in \mathcal{D}$ . La matrice  $\Delta$  devient alors

$$\Delta = n_L^2 K \begin{bmatrix} -2 & 1 & & 1 \\ 1 & -2 & 1 & \\ & & \ddots & \\ 1 & & 1 & -2 \end{bmatrix}. \quad (5.28)$$

On considère différentes hiérarchies correspondant à  $L \in \{0, \dots, 5\}$  avec la taille de la grille la plus fine fixée à  $n_L = 512$  quel que soit le choix de  $L$ . Les opérateurs de prolongation et de restriction  $\mathbf{P}_{\ell-1}^\ell$  et  $\mathbf{R}_\ell^{\ell-1}$  utilisés par le MLMC sont ceux définis dans les équations (3.19) et (3.20) et les opérateurs de filtrage  $\mathbf{S}_\ell$  utilisés par le F-MLMC sont les filtres de Shapiro d'ordre 2 définis dans l'équation (4.5). Comme pour les expériences numériques effectuées précédemment, on étudie deux longueurs de corrélation différentes. En fixant  $m = 10$  (et donc  $q = 5$ ), on peut définir les matrices  $\mathbf{A}_L^{006}$  et  $\mathbf{A}_L^{001}$  de taille  $512 \times 512$  à partir de l'équation (5.18), correspondant à une longueur de corrélation de  $D = 0.06 \approx 30n_L^{-1}$  et  $D = 0.01 \approx 5n_L^{-1}$ , respectivement. On s'intéresse aux problèmes d'estimation de  $\mathbf{v}^{006} := \mathbb{E}[(\mathbf{A}_L^{006} \mathbf{X}_L) \odot (\mathbf{A}_L^{006} \mathbf{X}_L)]$  et  $\mathbf{v}^{001} := \mathbb{E}[(\mathbf{A}_L^{001} \mathbf{X}_L) \odot (\mathbf{A}_L^{001} \mathbf{X}_L)]$  pour  $\mathbf{X}_L \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{I}_{n_L})$ . Le tenseur de diffusion étant homogène,  $\mathbf{v}^{006}$  et  $\mathbf{v}^{001}$  sont des vecteurs constants. On peut donc obtenir ces vecteur en calculant seulement une de leurs entrées en appliquant l'opérateur  $\mathbf{L}$  à un vecteur de la base canonique (cf. équation (5.4)). On trouve  $\mathbf{v}_i^{006} \approx 6.37$  et  $\mathbf{v}_i^{001} \approx 38.2$  pour tout  $i = 0, \dots, n_L - 1$ . Dans un tel contexte, utiliser un estimateur MC pour estimer les coefficients de normalisation n'a que peu d'intérêt en pratique, ici cela nous permet de tester le MC (et le MLMC ensuite) sur un cas relativement simple. L'estimateur MC de ces quantités est défini par l'équation (5.23).

Pour l'estimateur MLMC il faut définir les modèles utilisés sur les différents niveaux. Le modèle du niveau  $L$  peut être représenté à l'aide d'une fonction  $f_L$  définie

comme

$$\begin{aligned} f_L: \mathbb{R}^{n_L} &\rightarrow \mathbb{R}^{n_L} \\ \mathbf{x} &\mapsto (\mathbf{A}_L \mathbf{x}) \odot (\mathbf{A}_L \mathbf{x}). \end{aligned} \quad (5.29)$$

Les modèles de plus basse fidélité ne sont pas définis avec l'opérateur de Galerkin comme c'était le cas dans les illustrations des sections 3.4 et 4.3. En effet, les opérateurs de Galerkin doivent être stockés explicitement et sont donc très peu utilisés en pratique. Pour cette application on construit les opérateurs sur les grilles plus grossières par rediscrétisation de l'équation (5.10), basée sur la matrice  $\Delta$  de l'équation (5.28), ce qui permet de stocker ces opérateurs sous forme d'une application à des vecteurs. On note  $\mathbf{A}_\ell$  l'opérateur rediscrétisé sur la grille de taille  $n_\ell$  et on définit les modèles

$$\begin{aligned} f_\ell: \mathbb{R}^{n_\ell} &\rightarrow \mathbb{R}^{n_\ell} \\ \mathbf{x} &\mapsto (\mathbf{A}_\ell \mathbf{x}) \odot (\mathbf{A}_\ell \mathbf{x}) \end{aligned} \quad (5.30)$$

sur les différents niveaux de fidélité pour  $\ell = 0, \dots, L-1$ . La principale différence entre cette expérience numérique et les illustrations des sections 3.4 et 4.3 est que les modèles de la suite  $(f_\ell)_{\ell=0}^L$  utilisée ne sont pas linéaires. On définit également la suite des modèles  $(\tilde{f}_\ell)_{\ell=0}^L$ , utilisée par l'estimateur MLMC  $\hat{\mathbf{v}}_L^{\text{MLMC}}$  de  $\mathbf{v}$ , comme

$$\begin{aligned} \tilde{f}_\ell: \mathbb{R}^{n_L} &\rightarrow \mathbb{R}^{n_L} \\ \mathbf{x} &\mapsto \mathbf{P}_\ell^L f_\ell(\mathbf{R}_L^\ell \mathbf{x}), \end{aligned} \quad (5.31)$$

avec  $\mathbf{P}_L^L = \mathbf{R}_L^L = \mathbf{I}_{n_L}$ . L'estimateur F-MLMC utilise la suite de modèles  $(\bar{f}_\ell)_{\ell=0}^L$  où

$$\begin{aligned} \bar{f}_\ell: \mathbb{R}^{n_L} &\rightarrow \mathbb{R}^{n_L} \\ \mathbf{x} &\mapsto \bar{\mathbf{P}}_\ell^L f_\ell(\bar{\mathbf{R}}_L^\ell \mathbf{x}), \end{aligned} \quad (5.32)$$

avec  $\bar{\mathbf{P}}_\ell^L$  et  $\bar{\mathbf{R}}_L^\ell$  les opérateurs de transfert des équations (4.1) et (4.2) incluant les étapes de filtrage  $\mathbf{S}_\ell$ , et  $\bar{\mathbf{P}}_L^L = \bar{\mathbf{R}}_L^L = \mathbf{I}_{n_L}$ . Le coût de calcul moyen nécessaire à une évaluation de  $\tilde{f}_\ell$  et  $\bar{f}_\ell$  en  $\mathbf{X}_L$  est considéré identique et est noté  $\mathcal{C}_\ell$ . On suppose ce coût linéaire avec la taille de la grille et on a donc  $\mathcal{C}_\ell = \mathcal{O}(n_\ell)$  pour  $\ell = 0, \dots, L$ . Le budget total imposé est  $\eta = 100\mathcal{C}_L$  et l'allocation optimale  $(M_\ell)_{\ell=0}^L$  des estimateurs MLMC et F-MLMC est donnée par l'équation (3.108), où les quantités  $\mathcal{V}_\ell$  sont estimées à partir d'un échantillon pilote de taille 1000.

Les définitions des modèles  $\tilde{f}_\ell$  et  $\bar{f}_\ell$  des équations (5.31) et (5.32) ne sont pas les seules manières de définir des modèles de différentes fidélités dont les entrées et sorties sont des vecteurs de  $\mathbb{R}^{n_L}$ . On peut, par exemple, définir la suite de modèles  $(\tilde{\tilde{g}}_\ell)_{\ell=0}^L$  où

$$\begin{aligned} \tilde{\tilde{g}}_\ell: \mathbb{R}^{n_L} &\rightarrow \mathbb{R}^{n_L} \\ \mathbf{x} &\mapsto (\mathbf{P}_\ell^L \mathbf{A}_\ell \mathbf{R}_L^\ell \mathbf{x}) \odot (\mathbf{P}_\ell^L \mathbf{A}_\ell \mathbf{R}_L^\ell \mathbf{x}), \end{aligned} \quad (5.33)$$

et on pourrait, de la même manière, définir la suite  $(\bar{\bar{g}}_\ell)_{\ell=0}^L$  utilisant les opérateurs filtrés  $\bar{\bar{\mathbf{P}}}_\ell^L$  et  $\bar{\bar{\mathbf{R}}}_L^\ell$ . Ces modèles ont été testés sur le problème de normalisation mais ne sont pas présentés ici car les résultats obtenus sont extrêmement similaires à ceux obtenus avec les modèles  $\tilde{f}_\ell$  et  $\bar{f}_\ell$ . De plus, les modèles  $\tilde{\tilde{g}}_\ell$  ont un coût moyen d'application à  $\mathbf{x}$  légèrement supérieur à celui des modèles  $\tilde{f}_\ell$  à cause d'une application supplémentaire de  $\mathbf{P}_\ell^L$  et du produit de Schur sur des vecteurs de plus grande taille.

Il est important de noter que, contrairement à l'estimateur MC (5.23), l'estimateur

MLMC (3.7) ne permet pas de garantir la positivité des entrées des réalisations du vecteur  $\hat{\mathbf{v}}_L^{\text{MLMC}}$ , alors que  $\mathbf{v}$  est un vecteur de variances dont les entrées sont positives ou nulles. L'estimation MC sur le niveau grossier aboutit à un vecteur dont toutes les entrées sont positives ou nulles, mais l'ajout de termes de correction pouvant être négatifs produit un estimateur dont certaines réalisations peuvent avoir des entrées négatives. Ne pas pouvoir garantir la positivité de l'estimation est un problème connu de l'estimateur MLMC [48, 49] qui peut, selon l'application, être dommageable. Pour le problème de normalisation, cela veut dire que la matrice de corrélation  $\mathbf{C}$  de l'équation (5.1) reconstruite à partir d'une estimation MLMC de  $\mathbf{v}$  n'est pas forcément semi-définie positive. Cependant, aucune estimation négative n'a été obtenue sur toutes les expériences présentées ci-dessous. Pouvoir quantifier la probabilité d'obtenir une estimation négative pourrait permettre de mieux juger de l'intérêt du MLMC pour un tel problème.

### 5.2.1 Résultats du F-MLMC

On commence par présenter, sur la figure 5.1, les vecteurs  $\mathbf{v}^{006}$  et  $\mathbf{v}^{001}$  ainsi que des exemples d'estimation par les différents estimateurs utilisés.

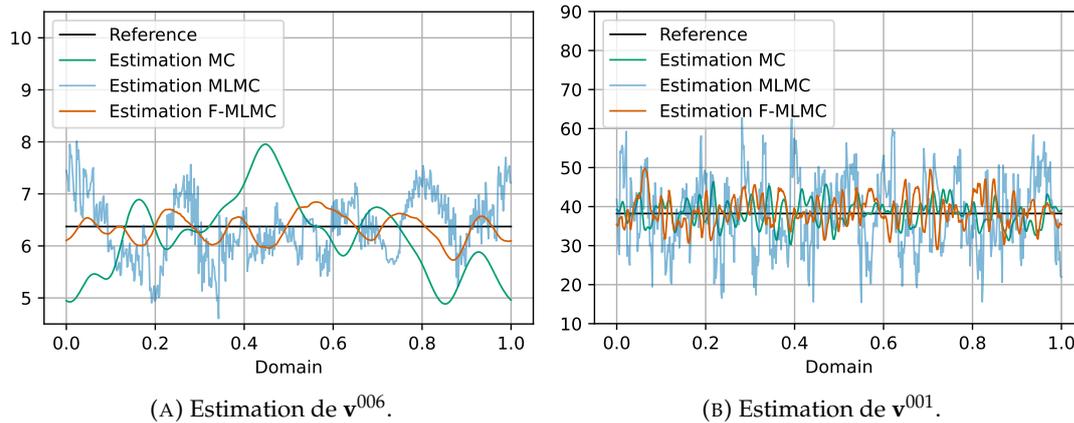


FIGURE 5.1 – Estimation de  $\mathbf{v}^{006}$  et  $\mathbf{v}^{001}$  par un MC, un MLMC et un F-MLMC à 6 niveaux ( $L = 5$ ) avec le budget total  $\eta = 100C_L$  et les allocations optimales présentées dans les figures 5.3a et 5.6a.

Malgré l'utilisation de modèles non linéaires, on observe, de manière similaire à la figure 4.3 de l'illustration 1D linéaire, symétrique et circulante, que l'estimation MLMC contient des composantes fines échelles (ou hautes fréquences) qui disparaissent dans l'estimation F-MLMC. Cette observation est particulièrement visible dans le cas de l'estimation de  $\mathbf{v}^{006}$ . Pour l'estimation de  $\mathbf{v}^{001}$ , le modèle fin  $f_L$ , utilisé notamment par l'estimateur MC, introduit déjà des composantes fines échelles qui peuvent se confondre avec celles que l'opérateur de prolongation introduit. En effet, la valeur de  $D = 0.01$  correspond, environ, à une corrélation sur 5 cellules, contrairement à la corrélation sur 30 cellules de  $D = 0.06$ . La figure 5.2 présente les valeurs absolues des espérances et variances des termes de correction,  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$ , définis dans l'équation (3.109), du MLMC et du F-MLMC ainsi que l'estimation des quantités  $\alpha$  et  $\beta$  du théorème 1, et cela pour le problème de l'estimation de  $\mathbf{v}^{006}$ .

Les résultats sont, encore une fois, très similaires à ceux présentés dans la figure 4.4 pour l'illustration 1D. Les relations supposées par le théorème 1, à savoir  $\mathcal{E}_\ell \lesssim n_\ell^{-\alpha}$  et  $\mathcal{V}_\ell \lesssim n_\ell^{-\beta}$ , semblent vérifiées pour  $\ell = 0, \dots, L$ . Les quantités  $\alpha$  et  $\beta$  sont plus élevées pour l'estimateur F-MLMC, ce qui indique que les espérances et variances des termes

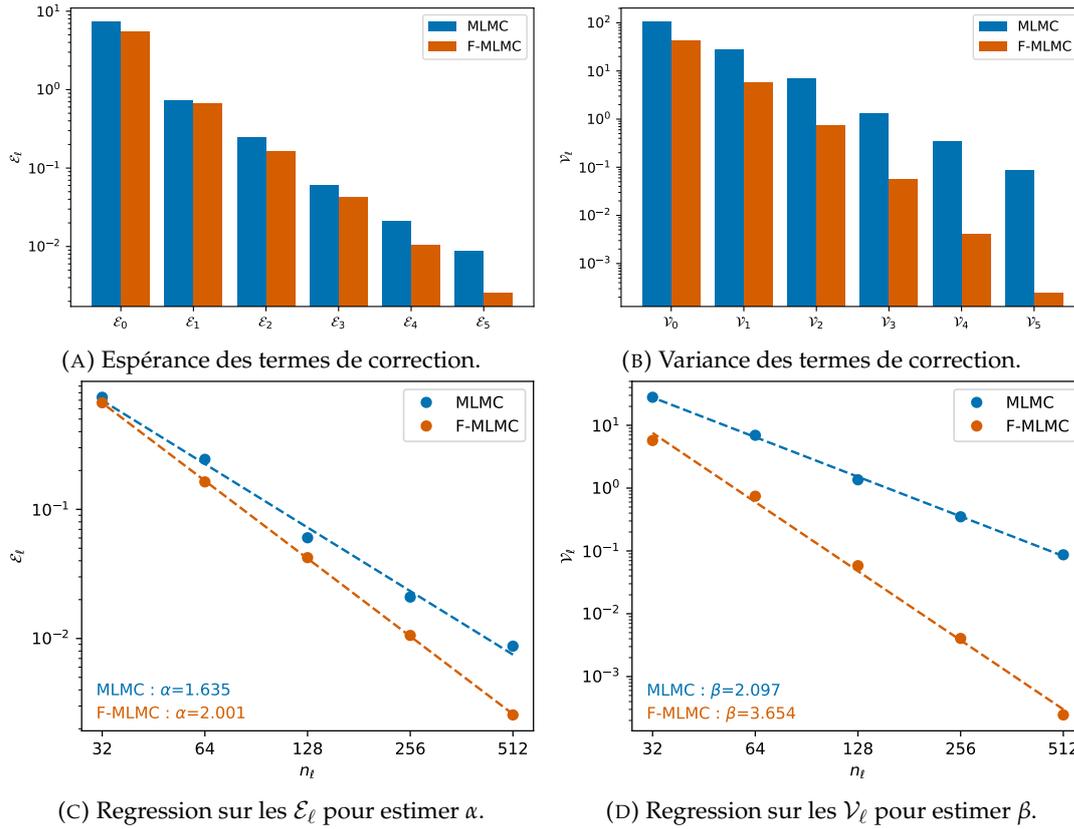


FIGURE 5.2 – Estimation des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  pour  $\ell = 0, \dots, L$  (haut) et regression sur ces quantites afin d'estimer  $\alpha$  et  $\beta$  (bas) pour le probleme d'estimation de  $\mathbf{v}^{006}$ . Les estimations sont faites a partir d'un echantillon de taille 1000.

de correction décroissent plus vite à travers les différents niveaux. En plus d'une décroissance plus rapide, on a des valeurs de  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  plus faibles pour le F-MLMC que pour le MLMC, et ce, pour tout  $\ell = 0, \dots, L$ . On s'attend donc à ce que les estimateurs F-MLMC puissent atteindre une variance plus faible que les estimateurs MLMC pour le même budget total  $\eta$ . C'est effectivement ce que l'on observe avec la figure 5.3 qui montre l'allocation optimale pour le MLMC et le F-MLMC à 6 niveaux, ainsi que les variances totales de ces estimateurs pour  $L \in \{0, \dots, 5\}$ .

Les estimateurs F-MLMC atteignent bien une variance inférieure aux estimateurs MLMC pour tout  $L \in \{1, \dots, 5\}$ , et ces estimateurs ont bien tous une variance plus faible que celle de l'estimateur MC. On observe cependant que pour le MLMC, la variance minimale est atteinte pour  $L = 3$ . L'ajout des niveaux les plus grossiers associés aux grilles de taille 32 et 16 détériorent l'estimateur MLMC, tandis que ces mêmes niveaux permettent au F-MLMC d'atteindre des variances encore plus faibles. Il est possible de mieux comprendre ces résultats en étudiant la décomposition de la variance de ces estimateurs dans un espace spectral. On réutilise ici la matrice de Hartley réordonnée  $\mathbf{H}_L^*$  de l'équation (3.23). Les variances spectrales et cumulées des estimateurs MLMC et F-MLMC sont affichées sur la figure 5.4.

Les différents résultats obtenus en conclusion de l'analyse spectrale théorique du MLMC et du F-MLMC pour des opérateurs linéaires, symétriques et circulants, se retrouvent, expérimentalement, malgré l'utilisation de modèles non-linéaires. En effet, on observe des profils de variance spectrale très similaires à ceux de la figure 4.6 de la section 4.3. La variance du MC provient surtout de l'estimation des grandes

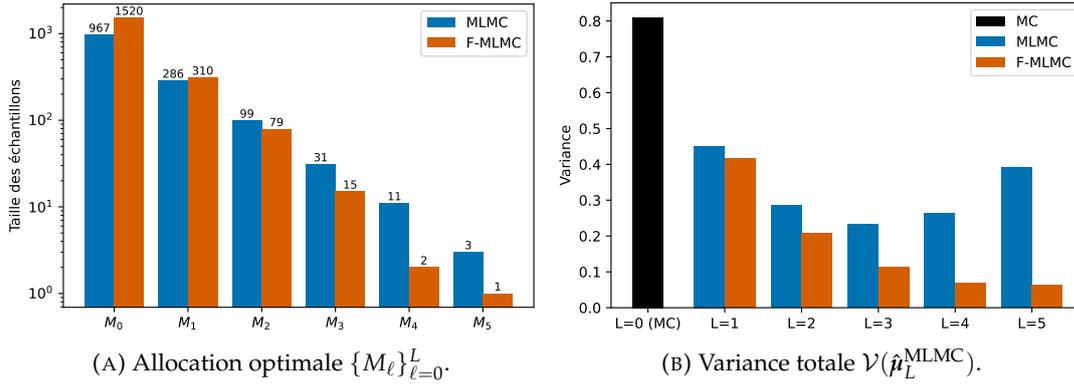
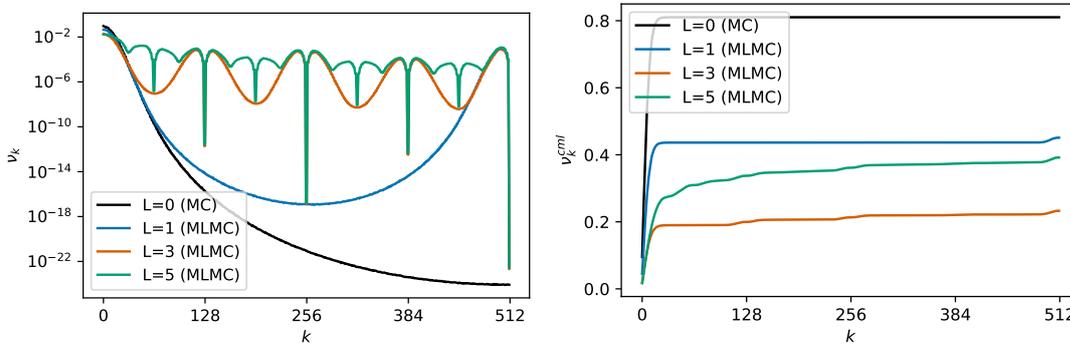
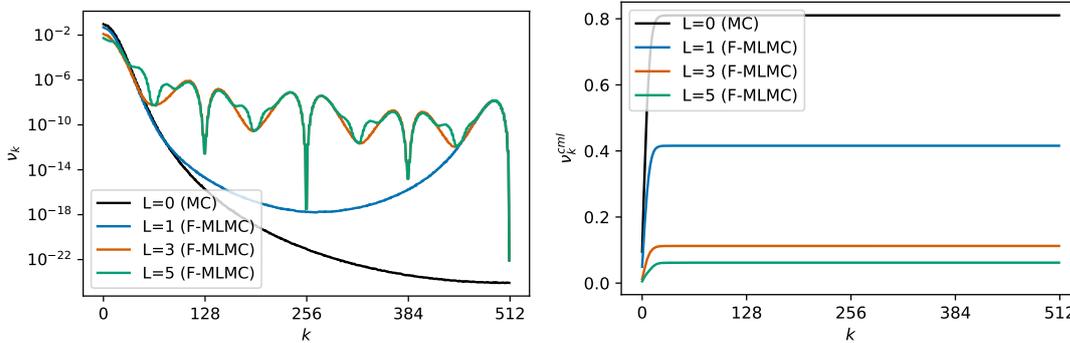


FIGURE 5.3 – Allocation optimale des échantillons pour les estimateurs MLMC et F-MLMC à 6 niveaux, et la variance totale résultante de l’estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC et F-MLMC pour l’estimation de  $\mathbf{v}^{006}$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100\mathcal{C}_L$ . La variance est estimée à partir de 1000 estimateurs.



(A) Variance spectrale  $\nu$  et variance cumulée  $\nu^{\text{cml}}$  des estimateurs MLMC.



(B) Variance spectrale  $\nu$  et variance cumulée  $\nu^{\text{cml}}$  des estimateurs F-MLMC.

FIGURE 5.4 – Variance spectrale et cumulée de l’estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC (haut) et F-MLMC (bas), avec  $L \in \{1, 3, 5\}$ , pour le problème d’estimation de  $\mathbf{v}^{006}$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100\mathcal{C}_L$ . La variance est estimée à partir de 1000 estimateurs.

échelles, le reste étant négligeable. Pour le MLMC, cependant, on observe les effets négatifs des fines échelles parasites introduites par les opérateurs de prolongation avec une variance importante sur ces échelles. Ces contributions sont assez faibles pour  $L = 1$  et sont cantonnées aux échelles les plus fines. Pour  $L = 3$  et  $L = 5$ , on

observe une détérioration de la variance qui s'étend à toutes les échelles et qui devient une partie non-négligeable de la variance totale. Les effets négatifs de la restriction se retrouvent dans l'estimation moins bonne des plus grandes échelles pour le MLMC à 6 niveaux que pour celui à 3 niveaux. L'ajout des filtres réduit fortement ces différents effets pour le F-MLMC, ce qui conduit à une variance plus faible pour l'ensemble des composantes, et donc une variance totale plus faible.

Les figures suivantes concernent le problème de l'estimation de  $\mathbf{v}^{001}$ . Pour ce problème, la matrice  $\mathbf{A}_L^{001}$  apparaissant dans le modèle fin  $f_L$  a un spectre plus étalé quand diagonalisée avec la matrice  $\mathbf{H}_L$  (cf. figure 3.9). Cela conduit à des champs discrétisés contenant plus de fines échelles, comme montré dans les exemples d'estimation de la figure 5.1. Ces fines échelles sont difficiles à estimer à l'aide des modèles plus grossiers et seront donc difficiles à estimer pour le MLMC. C'est ce qui est constaté sur l'histogramme des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  de la figure 5.5.

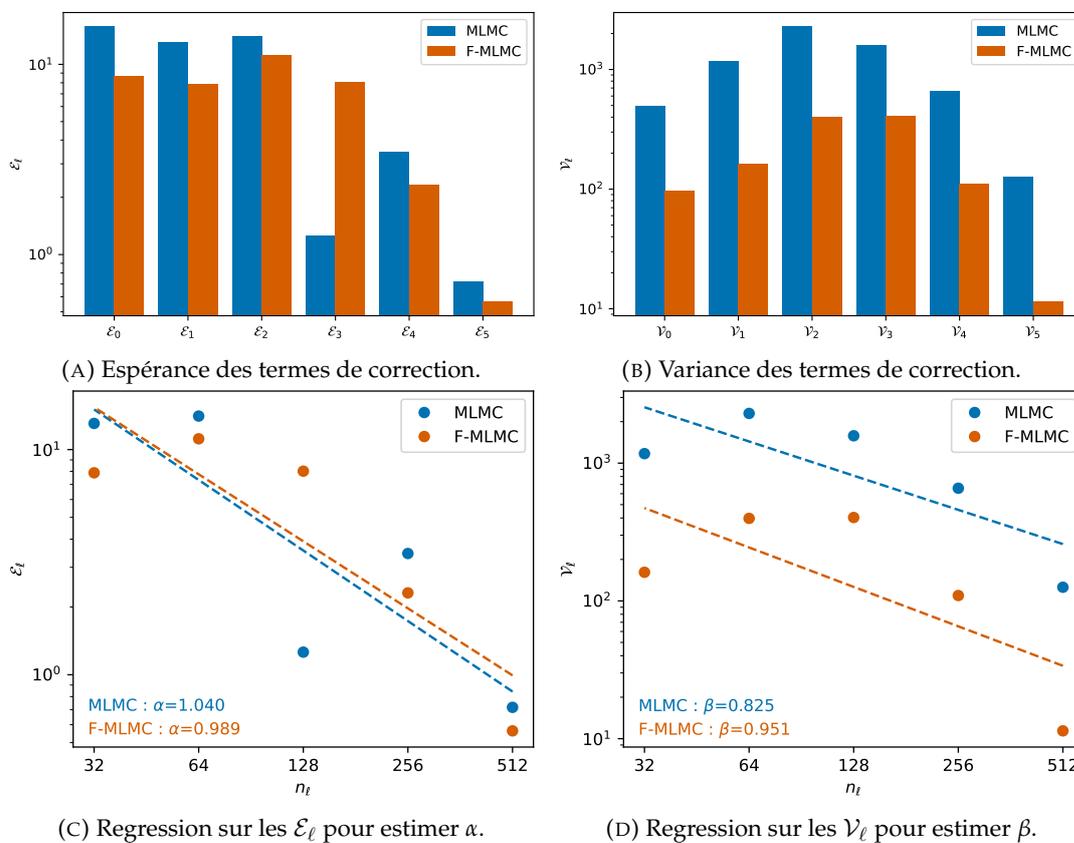


FIGURE 5.5 – Estimation des quantités  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  pour  $\ell = 0, \dots, L$  (haut) et regression sur ces quantités afin d'estimer  $\alpha$  et  $\beta$  (bas) pour le problème d'estimation de  $\mathbf{v}^{001}$ . Les estimations sont faites à partir d'un échantillon de taille 1000.

L'hypothèse de décroissance des espérances et variances des termes de correction à travers les 6 niveaux considérés n'est pas vérifiée ici, que ce soit pour le MLMC ou le F-MLMC. Cette décroissance n'est vérifiée que si l'on considère seulement les 2 ou 3 niveaux les plus fins. On observe, cependant, que le F-MLMC obtient constamment des valeurs plus faibles pour les  $\mathcal{E}_\ell$  et  $\mathcal{V}_\ell$  que le MLMC, indiquant qu'il est possible d'obtenir une variance plus faible pour cet estimateur. Les variances totales des estimateurs ainsi que l'allocation optimale pour les estimateurs à 6 niveaux sont présentés dans la figure 5.6.

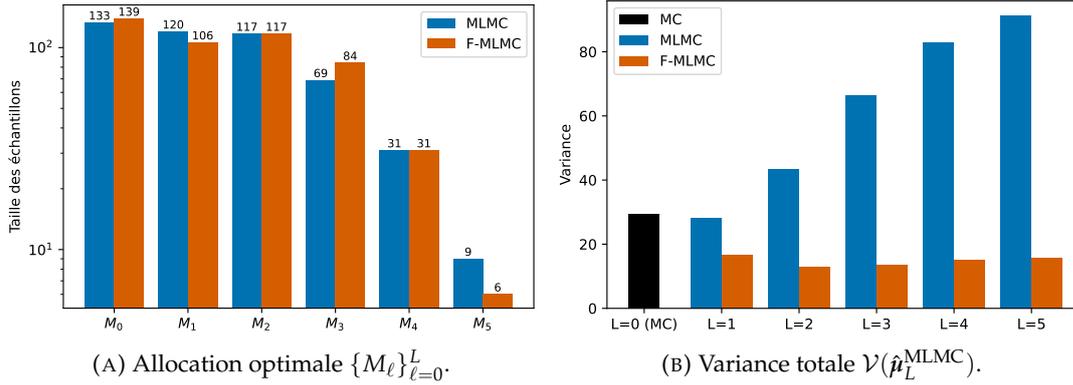


FIGURE 5.6 – Allocation optimale des échantillons pour les estimateurs MLMC et F-MLMC à 6 niveaux, et la variance totale résultante des estimateurs MC ( $L = 0$ ) et de différents estimateurs MLMC et F-MLMC pour le problème d'estimation de  $\mathbf{v}^{001}$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

Il est clair, en regardant l'allocation optimale pour le MLMC à 6 niveaux, qu'une partie bien plus importante du budget est consacrée aux ensembles des niveaux plus fins que pour l'estimation de  $\mathbf{v}^{006}$  de la figure 5.3, ce qui est beaucoup moins le cas pour le F-MLMC. Excepté le MLMC à 2 niveaux, tous les autres ont une variance totale bien plus élevée que celle du simple MC, l'utilisation de grilles plus grossières et les effets négatifs des opérateurs de transfert correspondants empirent l'estimateur. Ces phénomènes sont, encore une fois, fortement réduits par l'utilisation de filtres. L'ajout des niveaux les plus grossiers détériore également la variance totale du F-MLMC mais dans des proportions bien plus réduites, tous les estimateurs F-MLMC restant meilleurs que le MC. Les filtres semblent empêcher l'explosion de la variance lors de l'utilisation de niveaux non adaptés au problème considéré. Les variances spectrales et cumulées de la figure 5.7 confirment ces résultats.

Pour le MLMC, l'utilisation des niveaux les plus grossiers apporte des erreurs importantes sur l'entièreté du spectre. Le MLMC à 6 niveaux, par exemple, atteint une variance plus élevée que celle du MC pour toutes les composantes. Seul le MLMC à 2 niveaux parvient à une variance plus faible que celle du MC grâce à une meilleure estimation des grandes échelles. Le F-MLMC réduit fortement les contributions des moyennes et fines échelles à la variance totale, les rendant négligeables. L'estimation des grandes échelles est constamment meilleure que celle du MC, conduisant bien aux variances totales observées précédemment. Il faut noter, cependant, que les variances des grandes échelles se détériorent légèrement avec l'ajout des derniers niveaux. Les filtres de Shapiro d'ordre 2 semblent ne pas suffire pour ce cas. On peut alors tester d'autres opérateurs de filtrage afin de voir s'ils peuvent réduire encore plus la variance du F-MLMC. La figure 5.8 présente les variances totales d'estimateurs F-MLMC avec les filtres de Shapiro d'ordre 2,  $\mathbf{S}_\ell^{(2)}$ , et d'ordre 4,  $\mathbf{S}_\ell^{(4)}$ , ainsi que la double application de ces mêmes filtres,  $(\mathbf{S}_\ell^{(2)})^2$  et  $(\mathbf{S}_\ell^{(4)})^2$ , et ce, pour l'estimation de  $\mathbf{v}^{006}$  et  $\mathbf{v}^{001}$ . On rappelle que les spectres de ces filtres ont été présentés dans la figure 4.11.

Les variances totales obtenues avec les différents opérateurs sont relativement similaires. Pour l'estimation de  $\mathbf{v}^{006}$ , il est difficile de trouver un opérateur qui réduit la variance pour tout  $L$ , mais les filtres  $\mathbf{S}^{(4)}$  et  $(\mathbf{S}^{(4)})^2$  semblent légèrement meilleurs

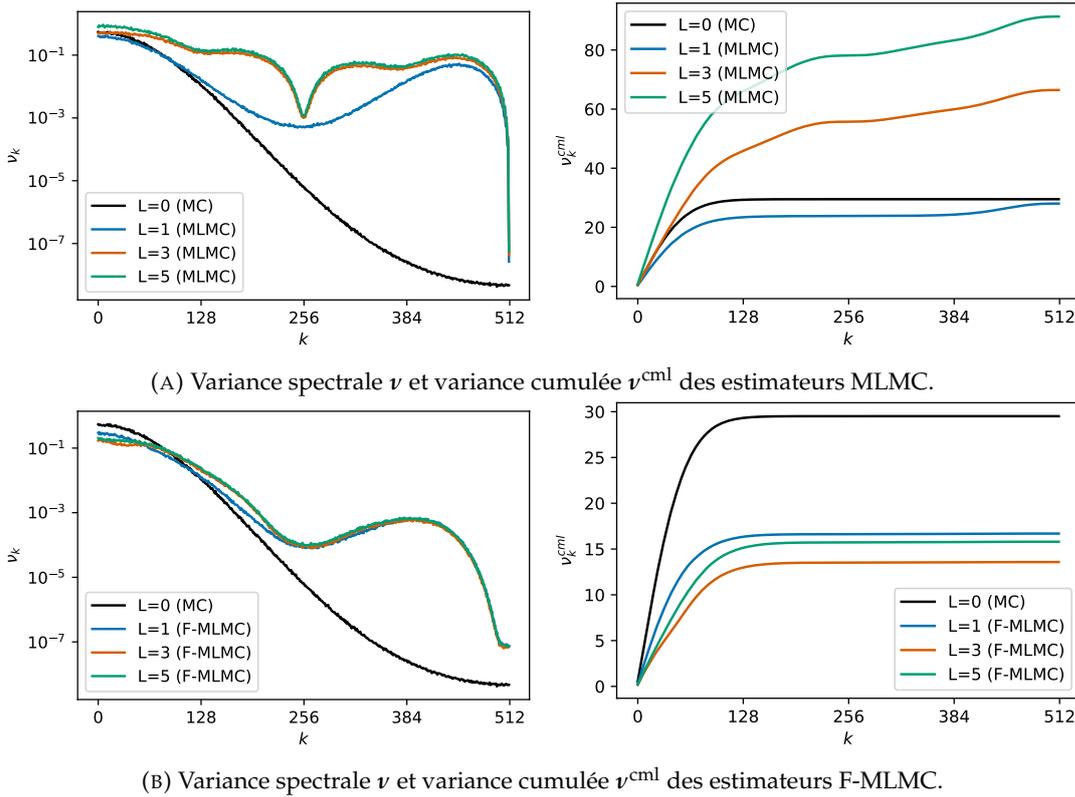


FIGURE 5.7 – Variance spectrale et cumulée de l’estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC (haut) et F-MLMC (bas), avec  $L \in \{1, 3, 5\}$ , pour le problème d’estimation de  $\mathbf{v}^{001}$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

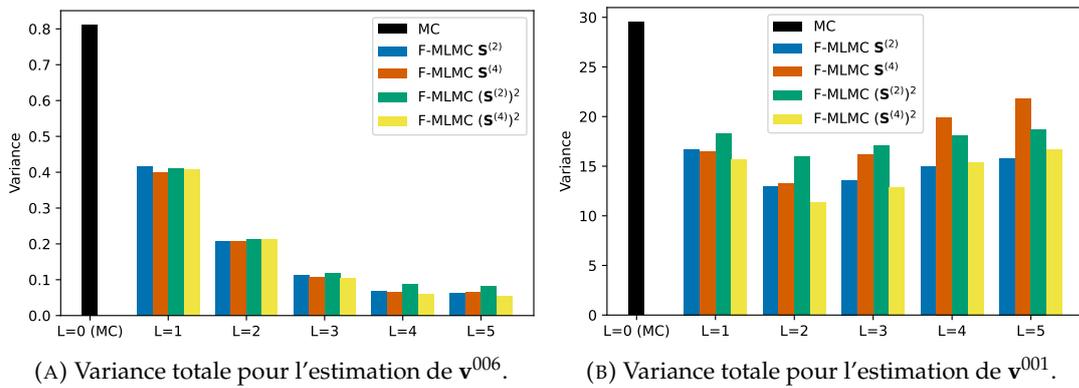


FIGURE 5.8 – Variance totale des estimateurs F-MLMC utilisant les filtres  $\mathbf{S}_\ell^{(2)}$  (bleu),  $\mathbf{S}_\ell^{(4)}$  (orange),  $(\mathbf{S}_\ell^{(2)})^2$  (vert) et  $(\mathbf{S}_\ell^{(4)})^2$  (jaune) pour l’estimation de  $\mathbf{v}^{006}$  (gauche) et de  $\mathbf{v}^{001}$  (droite). Les variances des estimateurs F-MLMC basés sur le filtre  $\mathbf{S}_\ell^{(2)}$  sont exactement les mêmes que précédemment et sont ici affichées pour faciliter la comparaison. Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 1000 estimateurs.

que les autres. Les différences sont plus visibles pour l’estimation de  $\mathbf{v}^{001}$ . Le filtre  $(\mathbf{S}_\ell^{(4)})^2$  permet d’atteindre des variances plus faibles que le filtre  $\mathbf{S}_\ell^{(2)}$  pour un nombre

modéré de niveaux ( $L < 4$ ). Pour  $L \geq 4$ , c'est  $\mathbf{S}^{(2)}$  qui atteint les valeurs de variance les plus faibles. Il reste important de constater que, par rapport aux variances totales des estimateurs MLMC des figures 5.3 et 5.6, l'ajout d'un filtre, quel qu'il soit parmi les quatre filtres testés ici, réduit la variance de manière considérable. Cependant, les gains apportés par le choix du filtre sont ici assez mineurs. Ces résultats, relativement similaires à ceux de la section 4.4.2 poussent, encore une fois, à effectuer une étude plus poussée sur l'ordre des opérateurs de transfert nécessaire pour minimiser la variance de l'estimateur. Il est possible qu'ici, la montée en ordre des opérateurs de transfert ne suffise pas à abaisser plus la variance de l'estimateur, ou, qu'au contraire, le filtre de Shapiro d'ordre 2 permette déjà d'atteindre une variance presque minimale.

## 5.2.2 Résultats du WMLMC

Sans forcément parler du choix des filtres, on aimerait vérifier que les opérateurs de transfert utilisés sont bien définis. Spécifiquement, et de la même manière que ce qui a été fait dans la section 4.4.3, on aimerait vérifier que, dans la relation  $\mathbf{R}_\ell^\ell = \alpha^{\ell-\ell}(\mathbf{P}_\ell^\ell)^\top$  que l'on a imposé, le choix de  $\alpha = 2^{-1/2}$  provenant de la propriété (4.46), est proche de celui qui minimise la variance totale de l'estimateur. Cette vérification était possible pour l'illustration 1D des sections précédentes car les modèles utilisés étaient linéaires. L'utilisation d'un estimateur WMLMC (4.51) permettait alors de trouver les poids  $\{\beta_\ell\}_{\ell=0}^{L-1}$  optimaux pour minimiser la variance, et ces poids pouvait être reliés à ceux du F-MLMC (4.50) pour retrouver la valeur optimale de  $\alpha$ . Ici, bien que les modèles  $\tilde{f}_\ell$  définis par les équations (5.30) et (5.32) ne soient pas linéaires, on a les relations

$$\tilde{f}_\ell(\mathbf{x}) = \tilde{\mathbf{P}}_\ell^L f_\ell(\tilde{\mathbf{R}}_L^\ell \mathbf{x}) \quad (5.34)$$

$$= \tilde{\mathbf{P}}_\ell^L [(\mathbf{A}_\ell \tilde{\mathbf{R}}_L^\ell \mathbf{x}) \odot (\mathbf{A}_\ell \tilde{\mathbf{R}}_L^\ell \mathbf{x})] \quad (5.35)$$

$$= \tilde{\mathbf{P}}_\ell^L [(\alpha^{L-\ell} \mathbf{A}_\ell (\tilde{\mathbf{P}}_\ell^L)^\top \mathbf{x}) \odot (\alpha^{L-\ell} \mathbf{A}_\ell (\tilde{\mathbf{P}}_\ell^L)^\top \mathbf{x})] \quad (5.36)$$

$$= \alpha^{2(L-\ell)} \tilde{\mathbf{P}}_\ell^L f_\ell((\tilde{\mathbf{P}}_\ell^L)^\top \mathbf{x}). \quad (5.37)$$

L'estimateur F-MLMC peut donc s'écrire, en notant  $\hat{f}_\ell: \mathbf{x} \mapsto \tilde{\mathbf{P}}_\ell^L f_\ell((\tilde{\mathbf{P}}_\ell^L)^\top \mathbf{x})$ ,

$$\hat{\mathbf{v}}_L^{\text{F-MLMC}} = \frac{1}{M_0} \sum_{i=1}^{M_0} \alpha^{2L} \hat{f}_0(\mathbf{X}_L^{(0,i)}) + \sum_{\ell=1}^L \frac{1}{M_\ell} \sum_{i=1}^{M_\ell} \alpha^{2(L-\ell)} \hat{f}_\ell(\mathbf{X}_L^{(\ell,i)}) - \alpha^{2(L-\ell+1)} \hat{f}_{\ell-1}(\mathbf{X}_L^{(\ell,i)}). \quad (5.38)$$

On utilise alors l'estimateur WMLMC défini dans l'équation (4.51) et la formule des poids  $\{\beta_\ell\}_{\ell=0}^{L-1}$  optimaux de l'équation (4.53) afin de vérifier que le coefficient  $\alpha = 2^{-1/2}$ , utilisé jusqu'à maintenant, est bien proche de la valeur optimale. Le tableau 5.1 présente les poids  $\{\beta_\ell\}_{\ell=0}^{L-1}$  obtenus pour le WMLMC basé sur les opérateurs de transfert sans filtres,  $\mathbf{P}_\ell^\ell$  et  $(\mathbf{P}_\ell^\ell)^\top$ , et les poids du F-WMLMC basé sur les opérateurs filtrés  $\tilde{\mathbf{P}}_\ell^\ell$  et  $(\tilde{\mathbf{P}}_\ell^\ell)^\top$ , appliqués à l'estimation de  $\mathbf{v}^{006}$ . Ces poids sont comparés aux poids  $\{2^{\ell-5}\}_{\ell=0}^5$  utilisés par le F-MLMC de l'équation (5.38).

Les poids obtenus par le WMLMC et le F-WMLMC sont relativement proches des poids utilisés par les estimateurs MLMC et F-MLMC. On observe une différence plus importante pour les poids des deux derniers niveaux,  $\ell = 0$  et  $\ell = 1$ , des estimateurs WMLMC avec 5 et 6 niveaux. Sans les filtres, les estimations sur ces niveaux sont fortement polluées par les composantes parasites introduites par les opérateurs de transfert, ce qui peut expliquer les différences entre les poids obtenus utilisés. Cela semble confirmé par les poids du F-WMLMC qui sont extrêmement semblables à

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$2^{\ell-5}$
0	0.4890	0.2342	0.1066	0.0399	0.0089	0.03125
1	1	0.4917	0.2378	0.1070	0.0373	0.0625
2		1	0.4938	0.2379	0.1059	0.125
3			1	0.4941	0.2374	0.25
4				1	0.4935	0.5
5					1	1

(A) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les WMLMC.

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur F-WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$2^{\ell-5}$
0	0.4981	0.2502	0.1261	0.0640	0.0318	0.03125
1	1	0.4993	0.2508	0.1265	0.0644	0.0625
2		1	0.4999	0.2508	0.1267	0.125
3			1	0.5002	0.2509	0.25
4				1	0.5002	0.5
5					1	1

(B) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les F-WMLMC.

TABLE 5.1 – Tables des poids optimaux  $\{\beta_\ell\}_{\ell=0}^L$  des estimateurs WMLMC et F-WMLMC pour le problème d'estimation de  $\mathbf{v}^{006}$  et pour différentes valeurs de  $L$ . Les poids sont obtenus à partir de l'équation (4.53) après résolution du problème (4.55) et avec estimation des covariances à partir d'un échantillon pilote de taille 10 000.

ceux utilisés par le F-MLMC, même sur les niveaux les plus grossiers. La figure 5.9 présente les variances totales de ces estimateurs.

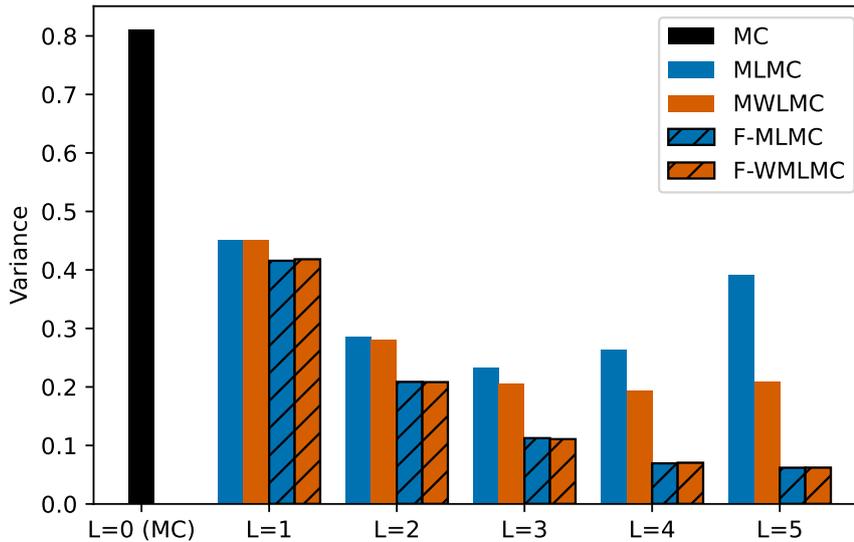


FIGURE 5.9 – Variance totale des estimateurs MLMC, WMLMC, F-MLMC et F-WMLMC pour  $L \in \{1, \dots, 5\}$  ainsi que l'estimateur MC ( $L = 0$ ) pour l'estimation de  $\mathbf{v}^{006}$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100\mathcal{C}_L$ . La variance est estimée à partir de 1000 estimateurs.

Sans surprise, l'obtention de poids similaires conduit à des variances totales quasiment identiques. C'est notamment le cas quand on compare le F-MLMC avec le F-WMLMC, ce qui confirme que le choix de  $\alpha = 2^{-1/2}$  est bien correct pour minimiser la variance des estimateurs. Pour le WMLMC, on observe que les différences entre les poids des niveaux les plus grossiers permet de fortement réduire la variance pour  $L = 4$  et  $L = 5$ . Les poids optimaux du WMLMC, qui sont plus faibles que ceux du MLMC, permettent d'utiliser les niveaux les plus grossiers sans trop détériorer la variance totale.

On reproduit maintenant les mêmes expériences pour le problème de l'estimation de  $\mathbf{v}^{001}$ . Les poids optimaux obtenus par les estimateurs WMLMC et F-WMLMC sont consignés dans le tableau 5.2.

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$2^{\ell-5}$
0	0.4136	0.1137	0.0182	0.0024	0.0001	0.03125
1	1	0.4058	0.0950	0.0081	0.0003	0.0625
2		1	0.3971	0.0883	0.0070	0.125
3			1	0.3948	0.0876	0.25
4				1	0.3943	0.5
5					1	1

(A) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les WMLMC.

Poids optimaux $\{\beta_\ell\}_{\ell=0}^L$ de l'estimateur F-WMLMC						
$\ell$	$L = 1$	$L = 2$	$L = 3$	$L = 4$	$L = 5$	$2^{\ell-5}$
0	0.5043	0.2554	0.1167	0.0400	0.0115	0.03125
1	1	0.5084	0.2527	0.1023	0.0268	0.0625
2		1	0.5077	0.2482	0.0966	0.125
3			1	0.5066	0.2468	0.25
4				1	0.5062	0.5
5					1	1

(B) Table des  $\{\beta_\ell\}_{\ell=0}^L$  pour les F-WMLMC.

TABLE 5.2 – Tables des poids optimaux  $\{\beta_\ell\}_{\ell=0}^L$  des estimateurs WMLMC et F-WMLMC pour le problème d'estimation de  $\mathbf{v}^{001}$  et pour différentes valeurs de  $L$ . Les poids sont obtenus à partir de l'équation (4.53) après résolution du problème (4.55) et avec estimation des covariances à partir d'un échantillon pilote de taille 10 000.

On observe, pour ce problème, une bien plus grande divergence entre les poids optimaux obtenus et ceux utilisés par le MLMC et le F-WMLMC. Les poids du niveau fin du WMLMC sont déjà plus faible que le facteur 0.5 utilisé par le MLMC, et les poids des niveaux les plus grossiers sont extrêmement faibles, descendant jusqu'à  $10^{-4}$  pour l'utilisation de la grille de taille 16 dans le MLMC à 6 niveaux. Les poids du FMLMC et du F-WMLMC sont bien plus semblables même si l'on observe des différences assez importantes sur les niveaux les plus grossiers. Malgré l'utilisation d'opérateurs de filtrage, les niveaux les plus grossiers semblent contenir assez de composantes parasites pour détériorer la variance totale. C'est ce que l'on peut observer sur la figure 5.10 qui présente les variance totales des différents estimateurs.

Les estimateurs WMLMC et F-WMLMC parviennent effectivement à atteindre des variances totales quasiment identiques ou plus faibles que celles des estimateurs

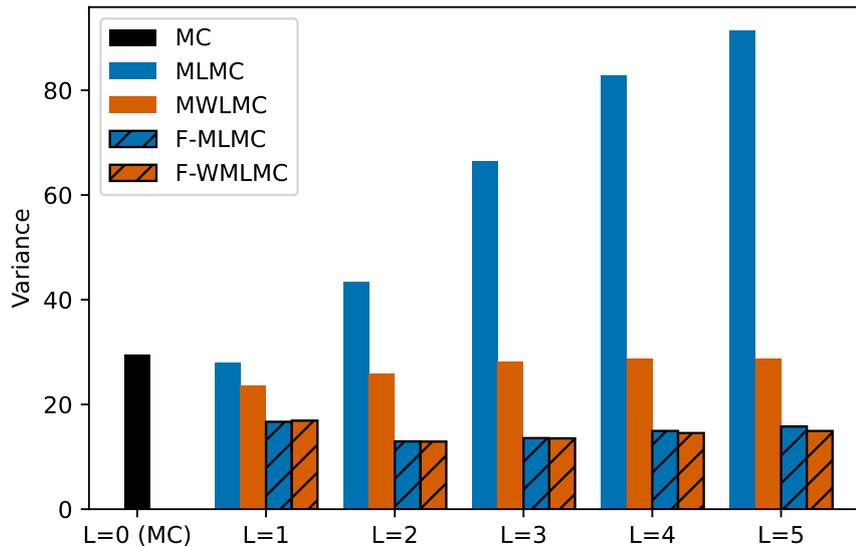


FIGURE 5.10 – Variance totale des estimateurs MLMC, WMLMC, F-MLMC et F-WMLMC pour  $L \in \{1, \dots, 5\}$  ainsi que l'estimateur MC ( $L = 0$ ) pour l'estimation de  $\mathbf{v}^{001}$ . Le niveau le plus fin  $L$  correspond toujours à une résolution avec  $n_L = 512$  et le budget total est fixé à  $\eta = 100\mathcal{C}_L$ . La variance est estimée à partir de 1000 estimateurs.

MLMC et F-MLMC. Pour le WMLMC, l'utilisation de poids bien plus faibles sur les niveaux grossiers permet d'empêcher la variance d'exploser en limitant fortement la contribution de ces niveaux. Les poids plus faibles pour les niveaux les plus grossiers ( $\ell = 0$  ou  $\ell = 1$ ) du F-WMLMC à 5 et 6 niveaux permettent également un léger gain de variance. L'estimateur à 3 niveaux reste, cependant, celui avec la plus faible variance. L'obtention, par le WMLMC, de poids plus faibles que ceux utilisés par le MLMC ou F-MLMC permet de réduire les effets négatifs dues à l'utilisation de niveaux trop grossiers pour le problème considéré.

### 5.2.3 Conclusion

Le problème de la normalisation en 1D a pour différence, comparé à l'illustration 1D utilisée dans les chapitres précédents, de se baser sur modèles non-linéaires. Cette non-linéarité des modèles ne permet donc pas d'appliquer la plupart des conclusions et résultats de l'analyse spectrale théorique développée pour le cas spécifique de modèles linéaires, symétriques et circulants. Cependant, les résultats des expériences numériques de cette section indiquent un comportement très similaire des estimateurs MLMC, F-MLMC et WMLMC pour l'estimation des coefficients de normalisation et pour l'illustration 1D des sections 3.4 et 4.3. Les opérateurs de transfert perturbent l'estimation de toutes les composantes dans des proportions qui peuvent être significatives pour la variance totale des estimateurs (cf. figure 5.4 par exemple). L'ajout des opérateurs de filtrage permet de réduire, en grande partie, ces effets négatifs, ce qui réduit la variance de toutes les composantes et donc la variance totale (cf. figures 5.3 et 5.6). Les tests des différents filtres permettent surtout de constater que l'ajout d'un filtre, quel qu'il soit, permet une réduction significative de la variance du MLMC. Le choix d'un filtre plutôt qu'un autre peut apporter quelques gains mineurs supplémentaires. L'estimateur WMLMC confirme que le coefficient  $\alpha = 2^{-1/2}$  utilisé dans la relation des opérateurs de restriction et de prolongation est

proche du coefficient optimal. Cet estimateur peut permettre, de plus, de réduire les effets négatifs d'une mauvaise utilisation des niveaux.

### 5.3 Normalisation 2D

Cette section traite du problème d'estimation des coefficients de normalisation dans un contexte 2D en considérant des champs tensoriels de diffusion  $\kappa$  hétérogènes. La principale différence lorsque le tenseur de diffusion  $\kappa$  est hétérogène en espace est que les champs de variance  $\mathbf{v}$  ne sont plus constants. Dans ce cas, il devient extrêmement coûteux de calculer les coefficients de normalisation exacts avec l'équation (5.4) car l'opérateur de diffusion  $\mathbf{L}$  doit être appliqué sur tous les vecteurs de la base canonique de  $\mathbb{R}^n$  avec  $n$  la taille de la grille considérée. C'est donc dans un tel contexte qu'il est souvent plus rentable, en terme de coût de calcul, d'estimer les coefficients de normalisation à l'aide d'un estimateur MC.

Le domaine considéré est  $\mathcal{D} = (0, 2) \times (0, 1) \subset \mathbb{R}^2$  avec des conditions aux bords périodiques dans les deux directions. Ce domaine est discrétisé à l'aide de  $L + 1$  grilles cartésiennes, centrées sur les cellules, de tailles  $n_\ell = n_\ell^x \times n_\ell^y$  pour  $\ell = 0, \dots, L$ . La grille la plus fine est de taille fixée  $n_L = 256 \times 128$ , et les tailles des autres grilles sont fixées par la relation  $n_\ell^x = 2^{\ell-L} n_L^x$  et  $n_\ell^y = 2^{\ell-L} n_L^y$ , ce qui donne donc  $n_\ell = 4^{\ell-L} n_L$  ainsi que  $n_\ell^x = 2 n_\ell^y$ . Les matrices de Gram associées aux différentes grilles sont  $\mathbf{W}_\ell = 2 n_\ell^{-1} \mathbf{I}_{n_\ell}$ . La figure 5.11 présente un exemple de ces grilles 2D centrées sur les cellules. On peut définir un opérateur  $\mathbf{A}_L \in \mathbb{R}^{n_L \times n_L}$  à partir de l'équation (5.18). Un tel opérateur est construit pour être appliqué à et pour retourner des vecteurs de taille  $n_L$ , dont les entrées sont associées aux centres des cellules de la grille, triées, premièrement par coordonnées croissantes dans la direction  $x$ , puis par coordonnées croissantes dans la direction  $y$ . Autrement dit, l'entrée d'un tel vecteur associée au centre d'une cellule située au point  $(x_i, y_j) \in \mathbb{R}^2$ , où  $x_i := 2(i + 1/2)/n_\ell^x$  et  $y_j := (j + 1/2)/n_\ell^y$ , est indexée par  $k = j n_\ell^x + i$ , pour  $i = 0, \dots, n_\ell^x - 1$  et  $j = 0, \dots, n_\ell^y - 1$ .

On commence par détailler la construction de la matrice  $\Delta \in \mathbb{R}^{n_L \times n_L}$  qui correspond à la discrétisation de l'opérateur différentiel  $\nabla \cdot \kappa \nabla$  de l'équation (5.8) sur la grille de taille  $n_L$  à l'aide des différences finies. On note, pour  $i = 0, \dots, n_L^x - 1$  et  $j = 0, \dots, n_L^y - 1$ ,  $x_{i,j}$  les points de l'espace qui correspondent au centre des cellules de la grille, c'est-à-dire les points représentés par des ● sur la figure 5.11. On note  $x_{i-\frac{1}{2},j}$  les points situés au centre des arêtes dans la direction  $i$ , ces points sont représentés par des ■, et  $x_{i,j-\frac{1}{2}}$  les points situés au centre des arêtes dans la direction  $j$ , représentés par des ◆. Par périodicité du domaine  $\mathcal{D}$ , on a  $x_{n_L^x-\frac{1}{2},j} = x_{-\frac{1}{2},j}$  pour tout  $j = 0, \dots, n_L^y - 1$ , et  $x_{i,n_L^y-\frac{1}{2}} = x_{i,-\frac{1}{2}}$  pour tout  $i = 0, \dots, n_L^x - 1$ .

On considère le vecteur  $\mathbf{u} = \{u_k\}_{k=0}^{n_L-1}$  qui représente un champ discrétisé au centre des cellules de cette grille et qui, en considérant la correspondance  $k = j n_L^x + i$  entre la numérotation des composantes du vecteur et la numérotation des cellules, est tel que  $u_{j n_L^x + i} = u_{i,j} = u(x_{i,j})$ . Le tenseur de diffusivité

$$\kappa = \begin{bmatrix} K_{11} & 0 \\ 0 & K_{22} \end{bmatrix}, \quad (5.39)$$

défini dans l'équation (5.9), est discrétisé aux centres des arêtes des cellules. Pour  $i = 0, \dots, n_L^x - 1$  et  $j = 0, \dots, n_L^y - 1$ , on note  $K_{i-\frac{1}{2},j}^{11} = K_{11}(x_{i-\frac{1}{2},j})$  et  $K_{i,j-\frac{1}{2}}^{22} = K_{22}(x_{i,j-\frac{1}{2}})$ .

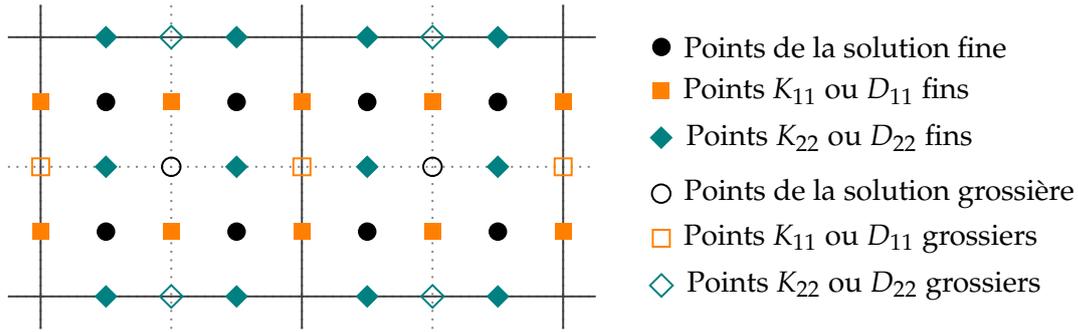


FIGURE 5.11 – Exemple d’une grille 2D centrée sur les cellules. Les ● représentent le centre des 8 cellules de la grille fine et ○ le centre des 2 cellules de la grille grossière. Les ◆ représentent le centre des arêtes de la grille fine dans la direction  $i$ , les ◇ ceux de la grille grossière. Les ■ représentent le centre des arêtes de la grille fine dans la direction  $j$ , les □ ceux de la grille grossière.

On pose, finalement,  $z(x) := \nabla \cdot \kappa \nabla u(x)$ ,  $\mathbf{z} = \{z_k\}_{k=0}^{n_L-1}$  le champ discrétisé correspondant qui est tel que  $z_j n_L^x + i = z_{i,j} = z(x_{i,j})$ , et  $\hat{\mathbf{z}}$  une approximation de  $\mathbf{z}$ . On utilise un schéma de différences finies centrées d’ordre 2 pour discrétiser  $\nabla \cdot \kappa \nabla$ , ce qui donne, pour  $i = 0, \dots, n_L^x - 1$  et  $j = 0, \dots, n_L^y - 1$

$$\hat{z}_{i,j} = (\Delta x)^{-2} \left( K_{i+\frac{1}{2},j}^{11} [u_{i+1,j} - u_{i,j}] - K_{i-\frac{1}{2},j}^{11} [u_{i,j} - u_{i-1,j}] \right) \quad (5.40)$$

$$+ (\Delta y)^{-2} \left( K_{i,j+\frac{1}{2}}^{22} [u_{i,j+1} - u_{i,j}] - K_{i,j-\frac{1}{2}}^{22} [u_{i,j} - u_{i,j-1}] \right) \quad (5.41)$$

$$= u_{i,j} (n_L^y)^2 \left( -K_{i-\frac{1}{2},j}^{11} - K_{i+\frac{1}{2},j}^{11} - K_{i,j-\frac{1}{2}}^{22} - K_{i,j+\frac{1}{2}}^{22} \right) \quad (5.42)$$

$$+ u_{i-1,j} (n_L^y)^2 K_{i-\frac{1}{2},j}^{11} + u_{i+1,j} (n_L^y)^2 K_{i+\frac{1}{2},j}^{11} \quad (5.43)$$

$$+ u_{i,j-1} (n_L^y)^2 K_{i,j-\frac{1}{2}}^{22} + u_{i,j+1} (n_L^y)^2 K_{i,j+\frac{1}{2}}^{22}, \quad (5.44)$$

avec  $\Delta x = (n_L^x/2)^{-1} = (n_L^y)^{-1}$  le pas de discrétisation en espace dans la direction  $x$  et  $\Delta y = (n_L^y)^{-1}$  le pas de discrétisation en espace dans la direction  $y$ . On note  $\gamma_{i,j} := K_{i-\frac{1}{2},j}^{11} + K_{i+\frac{1}{2},j}^{11} + K_{i,j-\frac{1}{2}}^{22} + K_{i,j+\frac{1}{2}}^{22}$  et on définit la matrice par blocs

$$\Delta = (n_L^y)^2 \begin{bmatrix} \mathbf{T}_0 & \mathbf{K}_1^{22} & & & \mathbf{K}_0^{22} \\ \mathbf{K}_1^{22} & \mathbf{T}_1 & \mathbf{K}_2^{22} & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \mathbf{K}_{n_L^y-1}^{22} \\ \mathbf{K}_0^{22} & & & \mathbf{K}_{n_L^y-1}^{22} & \mathbf{T}_{n_L^y-1} \end{bmatrix} \in \mathbb{R}^{n_L \times n_L} \quad (5.45)$$

où, pour  $j = 0, \dots, n_L^y - 1$ ,  $\mathbf{K}_j^{22} := \text{Diag}(\{K_{i,j-\frac{1}{2}}^{22}\}_{i=0}^{n_L^x-1}) \in \mathbb{R}^{n_L^x \times n_L^x}$  et

$$\mathbf{T}_j := \begin{bmatrix} -\gamma_{0,j} & K_{\frac{1}{2},j}^{11} & & & K_{-\frac{1}{2},j}^{11} \\ K_{\frac{1}{2},j}^{11} & -\gamma_{1,j} & K_{\frac{3}{2},j}^{11} & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & K_{n_L^x-\frac{3}{2},j}^{11} \\ K_{-\frac{1}{2},j}^{11} & & & K_{n_L^x-\frac{3}{2},j}^{11} & -\gamma_{n_L^x-1,j} \end{bmatrix} \in \mathbb{R}^{n_L^x \times n_L^x}. \quad (5.46)$$

La matrice  $\Delta$  est telle que  $\hat{\mathbf{z}} = \Delta \mathbf{u}$ .

On décrit maintenant la méthode utilisée pour générer le tenseur de diffusivité  $\kappa$  hétérogène défini dans l'équation (5.9) où  $m$  est fixé à 10 ( $q = 5$ ). On ne modélise pas directement les champs de diffusivité  $K_{11}$  et  $K_{22}$ , mais plutôt les deux champs  $D_{11}$  et  $D_{22}$  des longueurs de Daley. On modélise  $D_{11} = \zeta(\omega_1)$  et  $D_{22} = \zeta(\omega_2)$  comme deux réalisations d'un champ 2D gaussien  $\zeta$ , périodique sur le domaine  $\mathcal{D}$ . Ce champ aléatoire est de moyenne uniforme  $\mathbb{E}[\zeta(\mathbf{x})] = \mu_\zeta$ ,  $\forall \mathbf{x} \in \mathcal{D}$ , et a une structure de covariance quasi-gaussienne avec une variance uniforme  $\mathbb{V}[\zeta(\mathbf{x})] = \sigma_\zeta^2 = (\mu_\zeta/5)^2 = 0.04\mu_\zeta$ ,  $\forall \mathbf{x} \in \mathcal{D}$ , et une longueur de corrélation uniforme  $D_\zeta$ . Ces champs sont générés sous forme discrétisée au centre des arêtes des cellules de la grille de taille  $n_L = n_L^x \times n_L^y$ , et sont notés  $\mathbf{d}_{11}, \mathbf{d}_{22} \in \mathbb{R}^{n_L}$ . On a donc, pour  $i = 0, \dots, n_L^x - 1$  et  $j = 0, \dots, n_L^y - 1$ ,  $(\mathbf{d}_{11})_{jn_L^x+i} = D_{11}(x_{i-\frac{1}{2},j})$  et  $(\mathbf{d}_{22})_{jn_L^x+i} = D_{22}(x_{i,j-\frac{1}{2}})$ . En reprenant la figure 5.11 en exemple, le champ discrétisé  $\mathbf{d}_{11}$  (resp.  $\mathbf{d}_{22}$ ) correspond à la discrétisation de  $D_{11}$  (resp.  $D_{22}$ ) sur les points de l'espace représentés par les  $\blacksquare$  (resp.  $\blacklozenge$ ). Pour générer ces champs discrétisés  $\mathbf{d}_{11}$  et  $\mathbf{d}_{22}$  avec une structure de covariance quasi-gaussienne, on se base sur l'opérateur de diffusion en suivant l'équation (5.20). On utilise donc une matrice  $\mathbf{A}^{D_\zeta}$  définie dans l'équation (5.18) basée sur un champ tensoriel de Daley constant égal à  $D_\zeta$ . Par homogénéité du champ tensoriel de Daley, les coefficients de normalisation de l'opérateur  $\mathbf{L}^{D_\zeta} = \mathbf{A}^{D_\zeta}(\mathbf{A}^{D_\zeta})^T$  sont constants et égaux à  $\gamma_\zeta = (\mathbf{L}^{D_\zeta} \mathbf{e}_1)_1^{-1/2}$ . Après normalisation, on peut donc construire l'opérateur de variance uniforme  $\sigma_\zeta^2$  en multipliant simplement par  $\sigma_\zeta$ . Ainsi, en notant  $\mathbf{1}_{n_L} \in \mathbb{R}^{n_L}$  le vecteur constant égal à 1, on a, pour  $i = 1, 2$  et  $\mathbf{X} \sim \mathcal{N}(\mathbf{0}_{n_L}, \mathbf{I}_{n_L})$ ,

$$\mathbf{d}_{ii} = \mu_\zeta \mathbf{1}_{n_L} + \sigma_\zeta \gamma_\zeta \mathbf{A}^{D_\zeta} \mathbf{X}(\omega_i). \quad (5.47)$$

Il existe une probabilité que les champs  $\mathbf{d}_{11}$  et  $\mathbf{d}_{22}$  contiennent des valeurs négatives, ce qui poserait problème dans la définition du tenseur de diffusion  $\kappa$  et de la matrice  $\mathbf{A}_L$ . Un tel cas ne s'est pas produit dans les expériences présentées ici. Une fois les champs  $\mathbf{d}_{11}$  et  $\mathbf{d}_{22}$  générés, on obtient les tenseurs de diffusivité discrétisés avec la relation (5.9). Les expériences numériques qui suivent se basent sur deux tenseurs de diffusion, et donc deux couples  $(\mathbf{d}_{11}, \mathbf{d}_{22})$ . Le premier utilise les paramètres  $\mu_\zeta = 0.12$  et  $D_\zeta = 0.2$ , le second les paramètres  $\mu_\zeta = 0.02$  et  $D_\zeta = 0.04$ . Les réalisations de ces champs, qui sont utilisées dans les différentes expériences, sont présentées dans la figure 5.12.

À partir de ces champs on définit deux matrices  $\mathbf{A}_L^{012}$  et  $\mathbf{A}_L^{002}$  à partir de l'équation (5.18). Il est à noter que ces matrices ne sont pas circulantes dû à l'hétérogénéité de  $\kappa$ . On s'intéresse à l'estimation de  $\mathbf{v}^{012} := \mathbb{E}[(\mathbf{A}_L^{012} \mathbf{X}_L) \odot (\mathbf{A}_L^{012} \mathbf{X}_L)]$  et  $\mathbf{v}^{002} := \mathbb{E}[(\mathbf{A}_L^{002} \mathbf{X}_L) \odot (\mathbf{A}_L^{002} \mathbf{X}_L)]$ . Il est possible de profiter de la taille réduite du problème pour les calculer exactement à l'aide de l'équation (5.4). Ces vecteurs peuvent également être estimés avec l'estimateur MC de l'équation (5.23). Pour définir les

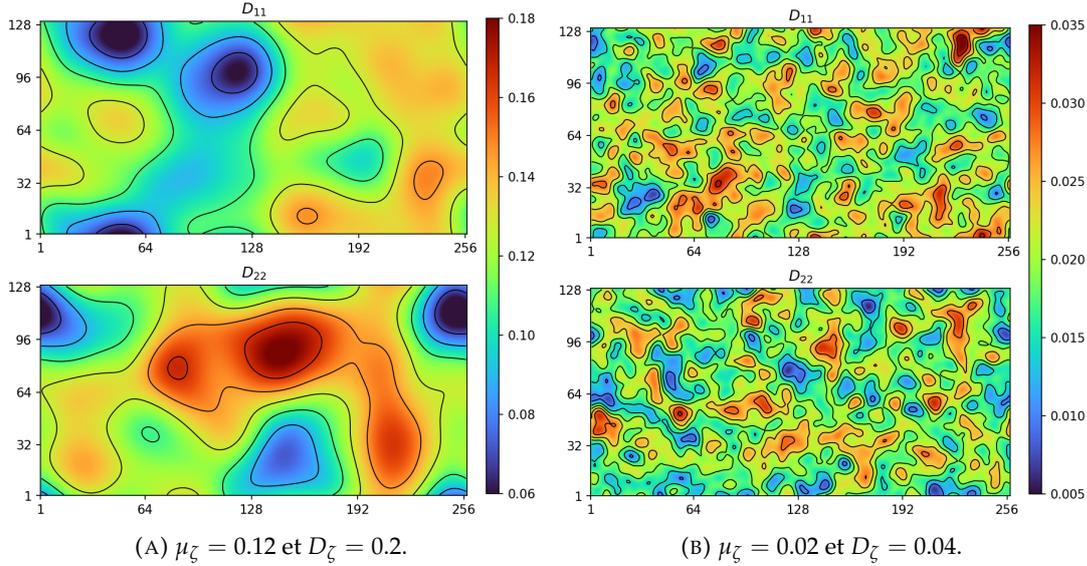


FIGURE 5.12 – Champs des longueurs de corrélation  $\mathbf{d}_{11}$  (haut) et  $\mathbf{d}_{22}$  (bas) utilisés pour définir le tenseur de diffusivité  $\kappa$ , avec  $\mu_\zeta = 0.12$  et  $D_\zeta = 0.2$  (gauche), et  $\mu_\zeta = 0.02$  et  $D_\zeta = 0.04$  (droite).

estimateurs MLMC et F-MLMC il faut définir la hiérarchie de modèles utilisée. On commence par définir les matrices  $\mathbf{A}_\ell^{012}$  et  $\mathbf{A}_\ell^{002}$  pour  $\ell = 0, \dots, L-1$  à partir de l'équation (5.18) sur la grille de taille  $n_\ell$ . Le tenseur de diffusion  $\kappa$  utilisé se base sur un tenseur de Daley construit à partir d'une rediscrétisation  $\mathbf{d}_{11}^\ell$  et  $\mathbf{d}_{22}^\ell$  des champs  $\mathbf{d}_{11}$  et  $\mathbf{d}_{22}$  par sélection, c'est-à-dire que l'on définit, pour  $i = 1, 2$ , par récurrence,

$$\mathbf{d}_{ii}^{\ell-1} := (\mathbf{R}_\ell^x \otimes \mathbf{R}_\ell^y) \mathbf{d}_{ii}^\ell \in \mathbb{R}^{n_{\ell-1}}, \quad (5.48)$$

où  $\otimes$  représente le produit de Kronecker entre deux matrices et où

$$\mathbf{R}_\ell^x := \begin{bmatrix} 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & & \\ & & & 1 & 0 \\ & & & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n_{\ell-1}^x \times n_\ell^x}, \quad (5.49)$$

$$\mathbf{R}_\ell^y := \begin{bmatrix} 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & & \\ & & & 1 & 0 \\ & & & & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n_{\ell-1}^y \times n_\ell^y}, \quad (5.50)$$

pour  $\ell = 1, \dots, L$ , sont des matrices de restriction 1D. Autrement dit, on construit  $\mathbf{d}_{11}^{L-1}$  et  $\mathbf{d}_{22}^{L-1}$  en ne prenant qu'une entrée sur quatre des champs  $\mathbf{d}_{11}^L = \mathbf{d}_{11}$  et  $\mathbf{d}_{22}^L = \mathbf{d}_{22}$ . Par récurrence, on peut construire des champs  $\mathbf{d}_{11}^\ell$  et  $\mathbf{d}_{22}^\ell$  pour tous les niveaux de fidélité plus grossiers  $\ell = 0, \dots, L-1$ . Ces restrictions des champs de longueur de Daley permettent bien de définir les matrices  $\mathbf{A}_\ell$  sur les différents niveaux de fidélité avec l'équation (5.10). On peut alors considérer la suite de modèles  $(f_\ell)_{\ell=0}^L$  de différentes fidélités définis suivant l'équation (5.30). Les opérateurs de prolongation et de restriction en 2D que l'on utilise pour le MLMC sont construits comme le produit de Kronecker des opérateurs 1D des équations (3.19) et (3.20). C'est-à-dire que, si l'on note  $(\mathbf{P}^x)_{\ell-1}^\ell \in \mathbb{R}^{n_\ell^x \times n_{\ell-1}^x}$  la prolongation 1D dans la direction  $x$  et  $(\mathbf{P}^y)_{\ell-1}^\ell \in \mathbb{R}^{n_\ell^y \times n_{\ell-1}^y}$

la prolongation 1D dans la direction  $y$ , on définit l'opérateur de prolongation 2D

$$\mathbf{P}_{\ell-1}^\ell := (\mathbf{P}^x)_{\ell-1}^\ell \otimes (\mathbf{P}^y)_{\ell-1}^\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}. \quad (5.51)$$

On a, de la même manière,

$$\mathbf{R}_\ell^{\ell-1} := (\mathbf{R}^x)_\ell^{\ell-1} \otimes (\mathbf{R}^y)_\ell^{\ell-1} \in \mathbb{R}^{n_{\ell-1} \times n_\ell}, \quad (5.52)$$

avec  $(\mathbf{R}^x)_\ell^{\ell-1}$  la restriction 1D dans la direction  $x$  et  $(\mathbf{R}^y)_\ell^{\ell-1}$  la restriction 1D dans la direction  $y$ , toutes deux définies à partir de l'équation (3.20). On retrouve bien, avec ces opérateurs de transfert 2D, les propriétés (3.15) et (3.16) de la section 3.2.2, i.e.,

$$(\mathbf{P}_{\ell-1}^\ell)^\top \mathbf{P}_{\ell-1}^\ell = \left( ((\mathbf{P}^x)_{\ell-1}^\ell)^\top \otimes ((\mathbf{P}^y)_{\ell-1}^\ell)^\top \right) \left( (\mathbf{P}^x)_{\ell-1}^\ell \otimes (\mathbf{P}^y)_{\ell-1}^\ell \right) \quad (5.53)$$

$$= \left( ((\mathbf{P}^x)_{\ell-1}^\ell)^\top (\mathbf{P}^x)_{\ell-1}^\ell \right) \otimes \left( ((\mathbf{P}^y)_{\ell-1}^\ell)^\top (\mathbf{P}^y)_{\ell-1}^\ell \right) \quad (5.54)$$

$$= 2\mathbf{I}_{n_{\ell-1}^x} \otimes 2\mathbf{I}_{n_{\ell-1}^y} \quad (5.55)$$

$$= 4\mathbf{I}_{n_{\ell-1}}, \quad (5.56)$$

$$\mathbf{R}_\ell^{\ell-1} = (\mathbf{R}^x)_\ell^{\ell-1} \otimes (\mathbf{R}^y)_\ell^{\ell-1} \quad (5.57)$$

$$= \frac{1}{\sqrt{2}} ((\mathbf{P}^x)_{\ell-1}^\ell)^\top \otimes \frac{1}{\sqrt{2}} ((\mathbf{P}^y)_{\ell-1}^\ell)^\top \quad (5.58)$$

$$= \frac{1}{2} (\mathbf{P}_{\ell-1}^\ell)^\top. \quad (5.59)$$

Le coefficient  $\alpha$  de l'équation (3.16) est donc égal à  $1/2$  en 2D. On en déduit que  $\mathbf{R}_\ell^{\ell-1} (\mathbf{R}_\ell^{\ell-1})^\top = \mathbf{I}_{n_{\ell-1}}$  et donc que

$$\mathbf{R}_\ell^{\ell-1} \mathbf{X}_\ell \sim \mathcal{N}(\mathbf{0}_{n_{\ell-1}}, \mathbf{I}_{n_{\ell-1}}), \quad \forall \mathbf{X}_\ell \sim \mathcal{N}(\mathbf{0}_{n_\ell}, \mathbf{I}_{n_\ell}), \quad (5.60)$$

pour  $\ell = 1, \dots, L$ . On retrouve bien la propriété (3.17) que l'on avait imposé en 1D dans le chapitre 3. En définissant les prolongations  $\mathbf{P}_\ell^L$  et les restrictions  $\mathbf{R}_L^\ell$  par composition suivant les équations (3.4) et (3.5), on peut définir la suite de modèles  $(\tilde{f}_\ell)_{\ell=0}^L$ , utilisés par l'estimateur MLMC, suivant l'équation (5.31) avec  $\mathbf{P}_L^L = \mathbf{R}_L^L = \mathbf{I}_{n_L}$ . Pour les modèles incluant l'étape de filtrage, on utilise le filtre de Shapiro d'ordre 2 en 2D construit par produit de Kronecker entre deux filtres de Shapiro d'ordre 2 en 1D,

$$\mathbf{S}_\ell := \mathbf{S}_\ell^x \otimes \mathbf{S}_\ell^y \in \mathbb{R}^{n_\ell \times n_\ell}, \quad (5.61)$$

où  $\mathbf{S}_\ell^x \in \mathbb{R}^{n_\ell^x \times n_\ell^x}$  et  $\mathbf{S}_\ell^y \in \mathbb{R}^{n_\ell^y \times n_\ell^y}$  sont les opérateurs de l'équation (4.5). On peut alors définir les opérateurs  $\tilde{\mathbf{P}}_L^\ell$  et  $\tilde{\mathbf{R}}_L^\ell$  à partir de l'équation (4.1) et de l'équation (4.2). La suite de modèles  $(\tilde{f}_\ell)_{\ell=0}^L$  utilisée par l'estimateur F-MLMC est définie par l'équation (5.32) avec  $\tilde{\mathbf{P}}_L^L = \tilde{\mathbf{R}}_L^L = \mathbf{I}_{n_L}$ . Le coût de calcul moyen nécessaire à une application de  $\tilde{f}_\ell$  et  $\tilde{f}_\ell$  à un vecteur aléatoire  $\mathbf{X}_\ell$  est considéré identique car les filtres  $\mathbf{S}_\ell$  sont des opérateurs très creux. On note un tel coût  $\mathcal{C}_\ell$ . On suppose ce coût linéaire avec la taille de la grille et on a donc  $\mathcal{C}_\ell = \mathcal{O}(n_\ell)$  pour  $\ell = 0, \dots, L$ . Le coût moyen d'une application de  $\tilde{f}_\ell$  est donc 4 fois supérieur à celui d'une évaluation de  $\tilde{f}_{\ell-1}$ . Le budget total imposé est  $\eta = 100\mathcal{C}_L$ , l'estimateur MC qui sert de comparaison se base donc sur un ensemble de taille 100 sur le niveau fin. Les allocations optimales  $\{M_\ell\}_{\ell=0}^L$  des estimateurs MLMC et F-MLMC sont calculées à partir de l'équation (3.108) où les quantités  $\mathcal{V}_\ell$  sont estimées à partir d'un échantillon pilote de taille 1000.

### 5.3.1 Résultats du F-MLMC

On commence par s'intéresser au problème de l'estimation de  $\mathbf{v}^{012}$ , utilisant donc le couple  $(\mathbf{d}_{11}, \mathbf{d}_{22})$  représenté sur la figure 5.12a, basé sur les paramètres  $\mu_\zeta = 0.12$  et  $D_\zeta = 0.2$ . La figure 5.13 présente la référence, c'est-à-dire la valeur exacte de  $\mathbf{v}^{012}$  calculée explicitement à partir de l'équation (5.4), ainsi qu'un exemple d'estimation MC, MLMC avec 2 et 4 niveaux, et F-MLMC avec 2 et 4 niveaux.

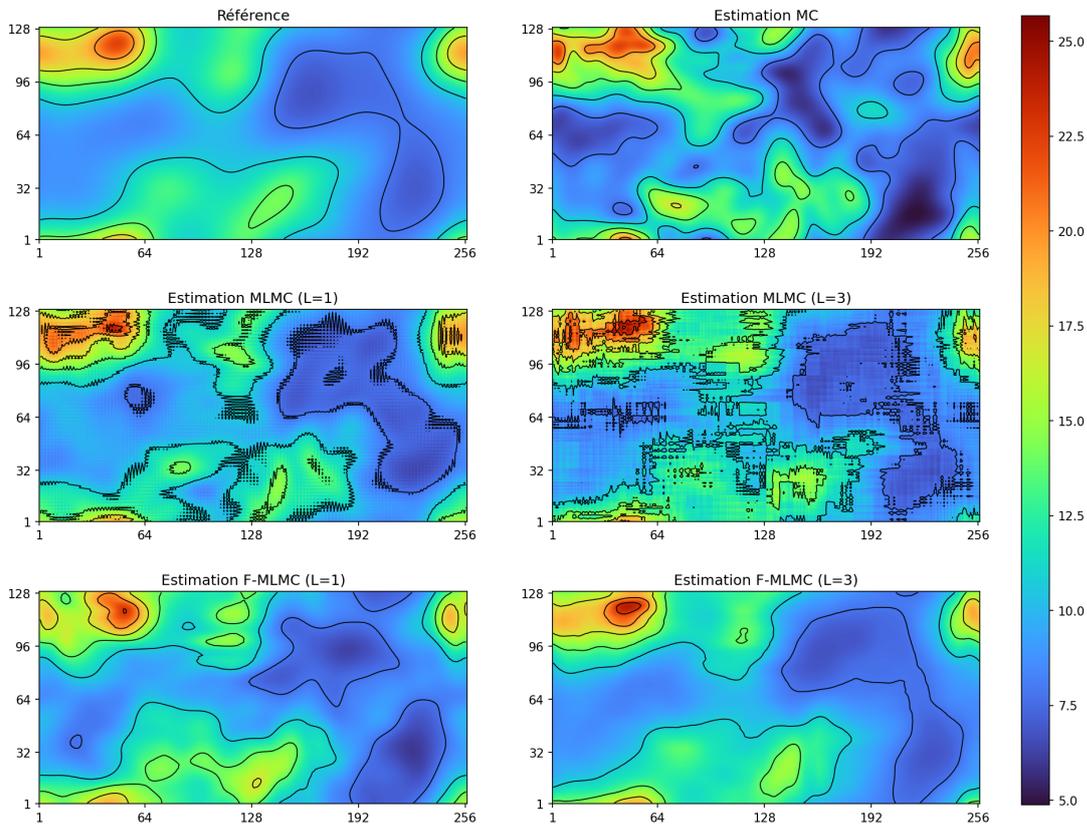


FIGURE 5.13 – Exemple d'estimations de  $\mathbf{v}^{012}$  par un MLMC et un F-MLMC à 2 et 4 niveaux avec le budget total  $\eta = 100C_L$ , comparées à la référence  $\mathbf{v}^{012}$ , en haut à gauche, et à une estimation par MC pour le même budget  $\eta$ , en haut à droite.

La principale observation ici est la présence de fines échelles (ou hautes fréquences) sur les estimations MLMC. Ces composantes semblent plus importantes pour le MLMC à 4 niveaux et contrastent avec les estimations F-MLMC qui sont, elles, bien plus lisses. On visualise clairement l'introduction de fines échelles parasites dues à l'opérateur de prolongation pour le MLMC, et leur atténuation lorsque l'on ajoute un opérateur de filtrage. On rappelle, cependant, que les différents estimateurs présentés ici sont tous non-biaisés et leur espérance est bien égale à  $\mathbf{v}^{012}$  malgré les interférences visuelles qui apparaissent sur les estimations MLMC de la figure 5.13. Cela est illustré avec la figure 5.14 qui affiche une estimation de l'espérance du MC, MLMC et F-MLMC à partir de 100 estimateurs.

On observe bien la caractéristique non-biaisée des différents estimateurs, à des erreurs d'échantillonnage près sur la moyenne de l'estimateur. Ils estiment bien tous, en moyenne, le champ  $\mathbf{v}^{012}$ , ce qui confirme donc que, dans un tel cas, l'erreur quadratique moyenne de ces estimateurs est composée uniquement de leur variance. C'est sur ce point que l'on s'attend à observer des différences entre les estimateurs. On peut calculer la variance des estimateurs en chaque point de la grille pour obtenir

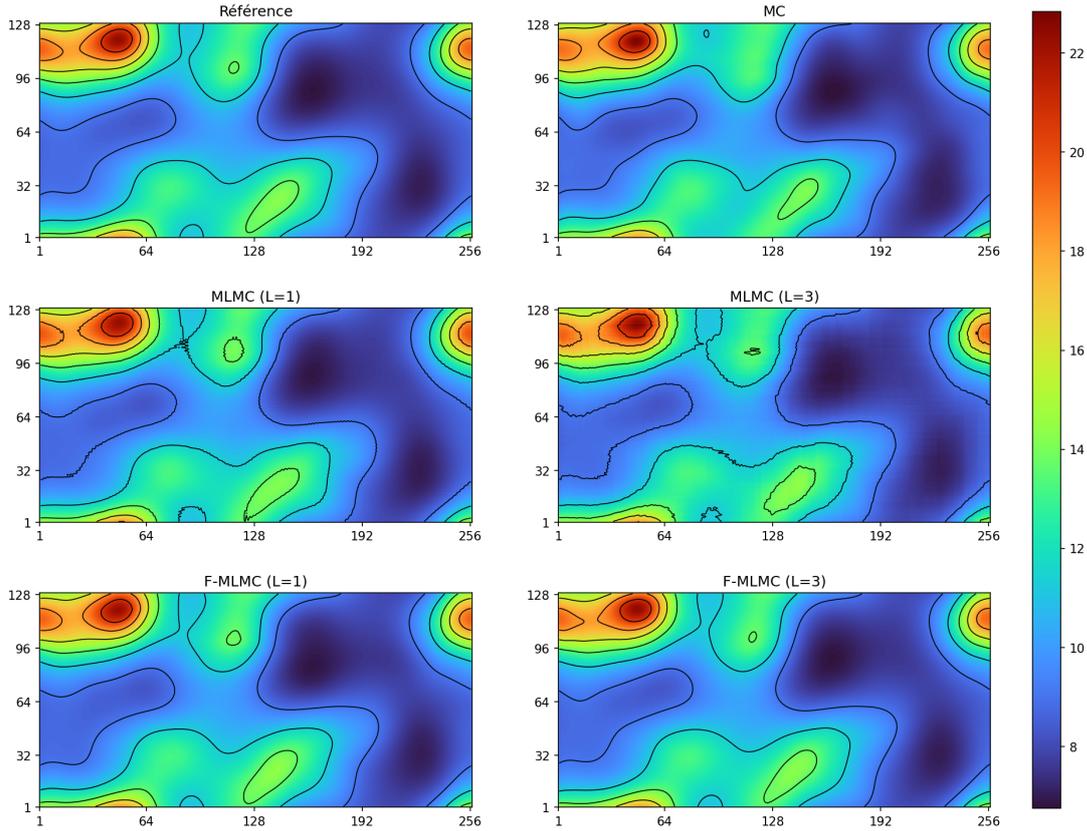


FIGURE 5.14 – Espérance des estimateurs MLMC et F-MLMC avec 2 et 4 niveaux, comparés à la référence  $\mathbf{v}^{012}$  et à l’espérance de l’estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . L’espérance est estimée à partir de 100 estimateurs.

le champ de variance de chacun d’eux. Ces champs sont présentés sur la figure 5.15.

Visuellement, le champ de variance du MLMC, notamment à 4 niveaux, est de norme inférieure à celle du MC. Cependant, on y retrouve clairement les fines échelles parasites qui apparaissaient dans la figure 5.13. Ces erreurs ont des conséquences importantes et visibles sur les estimations individuelles de  $\mathbf{v}^{012}$  comme démontré avec les exemples d’estimation de la figure 5.13. Le F-MLMC, lui aussi, possède un champ de variance de norme inférieure à celui du MC, mais sans les interférences fines échelles. Les filtres permettent de réduire fortement ces composantes dans la variance du F-MLMC. On observe également que F-MLMC à 4 niveaux a un champ de variance dont la norme est inférieure à celle des estimateurs MLMC, la variance totale a été réduite par l’ajout des filtres. Afin de mieux visualiser ces résultats, on affiche la décomposition spectrale de ces champs sur la figure 5.16. La décomposition spectrale est calculée à partir de la matrice de Hartley  $\mathbf{H}_L \in \mathbb{R}^{n_L \times n_L}$  pour projeter les champs de variance 2D dans l’espace spectral de Hartley. Cette matrice est construite par produit de Kronecker entre deux matrices de Hartley 1D  $\mathbf{H}_L^x \in \mathbb{R}^{n_L^x \times n_L^x}$  et  $\mathbf{H}_L^y \in \mathbb{R}^{n_L^y \times n_L^y}$  définies à partir de l’équation (3.21). La variance spectrale 2D est définie comme  $\mathbf{v} = (v_k)_{k=0}^{n_L-1} \in \mathbb{R}^{n_L}$  avec

$$v_k := 2n_L^{-1} \mathbb{E}[(\mathbf{h}_k^L)^T (\hat{\mathbf{v}}_L^{\text{MLMC}} - \mathbf{v})^2], \quad k = 0, \dots, n_L - 1, \quad (5.62)$$

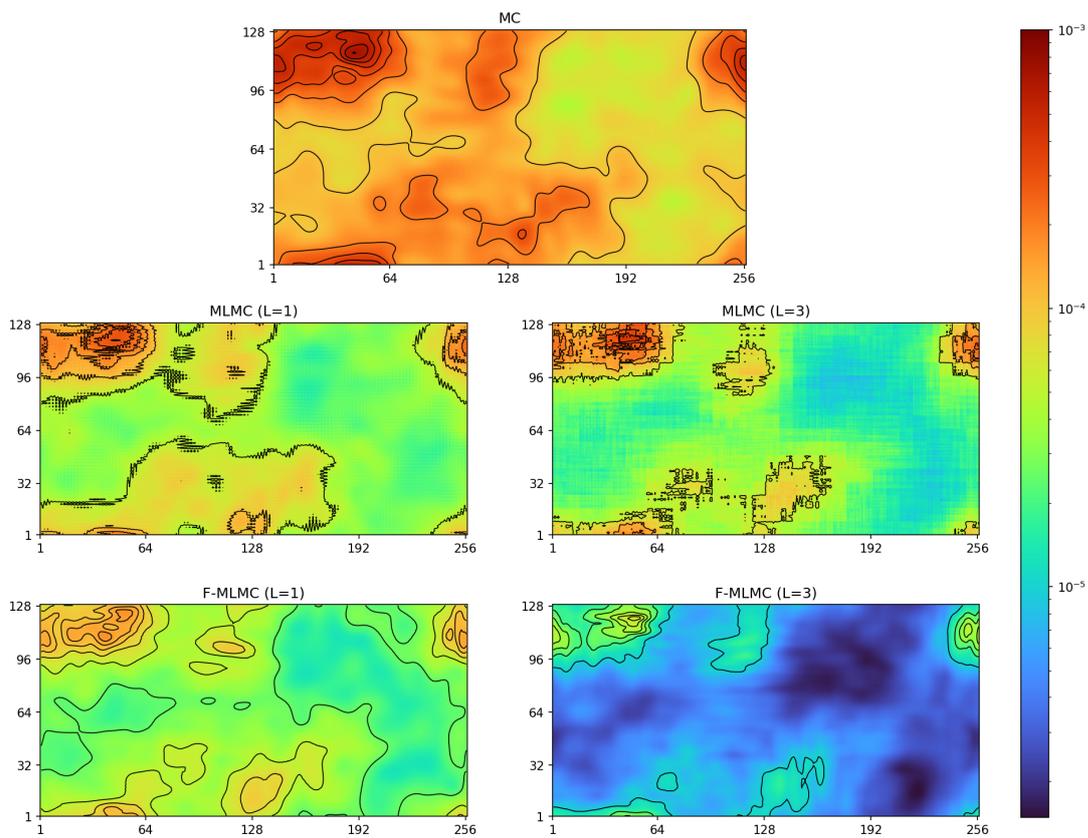


FIGURE 5.15 – Champs de variance des estimateurs MLMC et F-MLMC avec 2 et 4 niveaux de  $v^{012}$  comparés à celui de l'estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.

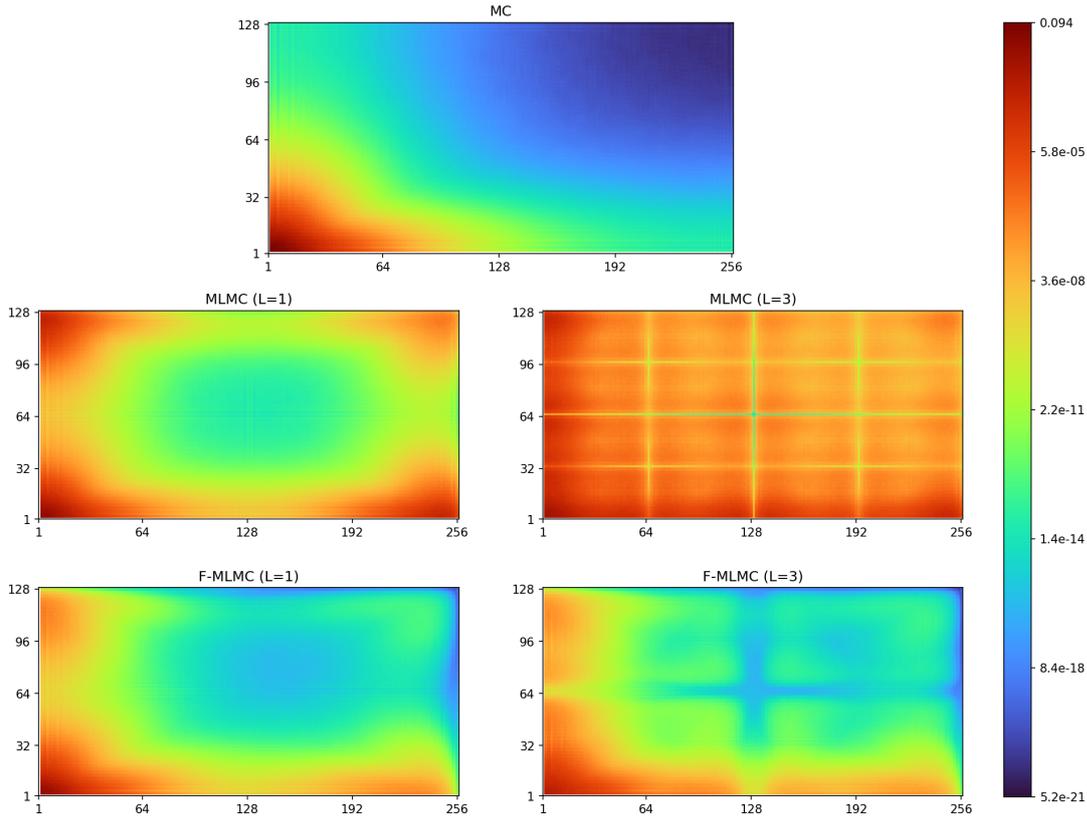


FIGURE 5.16 – Champs de variance spectrale des estimateurs MLMC et F-MLMC avec 2 et 4 niveaux de  $\mathbf{v}^{012}$  comparés à celui de l’estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.

où  $\mathbf{h}_k^L$  est la  $k$ -ième colonne de la matrice  $\mathbf{H}_L$ . Pour les différentes figures, les colonnes des matrices 1D  $\mathbf{H}_L^x$  et  $\mathbf{H}_L^y$  sont réordonnées selon l’équation (3.23). Le coin inférieur-gauche des champs présentés correspond donc aux grandes échelles tandis que le coin supérieur-droit correspond aux fines échelles. Les coins supérieur-gauche et inférieur-droit représentent les fines échelles dans une direction et les grandes échelles dans l’autre.

On retrouve, pour l’estimateur MC, un profil des variances spectrales similaire à ceux des problèmes 1D étudiés, par exemple la figure 5.4. En effet, on y observe une variance élevée pour les composantes grandes échelles et une variance très faible pour les composantes fines échelles. Le MLMC à deux niveaux (à gauche, ligne du milieu) présente, lui, une variance élevée sur les grandes et les fines échelles. Cette variance élevée s’étend à toutes les composantes pour le MLMC à quatre niveaux. On retrouve donc bien les composantes fines échelles introduites par les opérateurs de transfert qui apportent des erreurs sur l’ensemble des composantes, ces résultats sont similaires aux variances spectrales obtenues sur la figure 5.4. Finalement, on observe que l’ajout des filtres avec l’estimateur F-MLMC permet bien de réduire la variance des composantes fines échelles. Sur les figures des variances spectrales en 1D, on observait également une réduction importante de la variance sur les grandes échelles pour les estimateurs MLMC et F-MLMC, ce qui n’est pas immédiatement visible sur la figure 5.16. On construit donc un vecteur de variance cumulée 2D,  $\mathbf{v}^{\text{cml}} = (v_k^{\text{cml}})_{k=0}^{n_L^y-1} \in \mathbb{R}^{n_L^y}$ , calculé en sommant les entrées d’un champ de variance

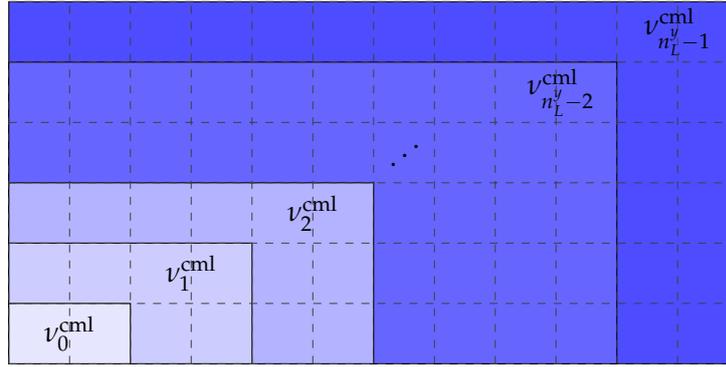
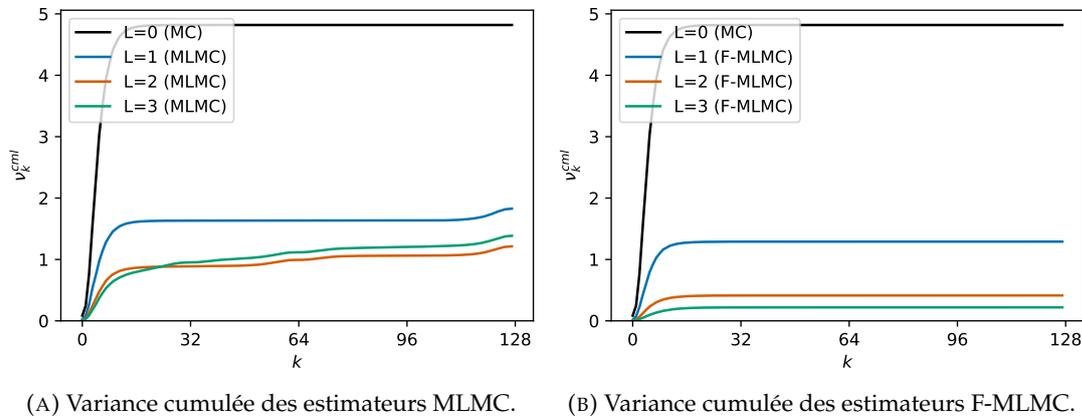


FIGURE 5.17 – Visualisation des entrées du vecteur des variances cumulées  $\nu^{\text{cml}}$ . L'entrée  $k$  du vecteur correspond à la somme des composantes du champ des variances spectrales dans le rectangle associé.



(A) Variance cumulée des estimateurs MLMC. (B) Variance cumulée des estimateurs F-MLMC.

FIGURE 5.18 – Variance cumulée  $\nu^{\text{cml}}$  des estimateurs MLMC (gauche) et F-MLMC (droite) avec  $L \in \{1, 2, 3\}$ , pour le problème d'estimation de  $\mathbf{v}^{012}$ , comparés à celle de l'estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.

spectral (5.62), tel que ceux de la figure 5.16, en motif rectangulaire à partir du coin inférieur-gauche, suivant le schéma de la figure 5.17. Ainsi, la première entrée du vecteur  $\nu^{\text{cml}}$  correspond à la somme des valeurs de variance spectrale des deux cellules situées sur la première ligne et deux premières colonnes de la grille, la seconde entrée de  $\nu^{\text{cml}}$  correspond à la somme sur les deux premières lignes et quatre premières colonnes de la grille, etc. On a donc

$$\nu_k^{\text{cml}} = \sum_{j=0}^k \sum_{i=0}^{2k+1} \nu_{jn_L^y+i}, \quad k = 0, \dots, n_L^y - 1. \quad (5.63)$$

Les premières composantes de  $\nu^{\text{cml}}$  correspondent à la variance cumulée des composantes grandes échelles, tandis que la dernière composante  $\nu_{n_L^y-1}^{\text{cml}}$  correspond à la variance totale de l'estimateur. Les variances cumulées des différents estimateurs sont présentées sur la figure 5.18.

On y retrouve, encore une fois, des profils très similaires à ceux obtenus pour le problème 1D sur la figure 5.4. Pour l'estimateur MC, la quasi-totalité de la variance provient de l'estimation des grandes échelles, le reste étant négligeable. Les estimateurs MLMC, bien qu'ayant une variance supérieure sur l'ensemble des composantes

moyennes et fines échelles, parviennent à fortement réduire la variance des grandes échelles, aboutissant ainsi à une variance totale plus faible que le MC. Les erreurs apportées par les opérateurs de transfert sont visibles mais les améliorations sur l'estimation des grandes échelles les compensent largement. Les variances cumulées du F-MLMC permettent de constater l'amélioration de l'estimation de toutes les composantes, et donc une réduction de la variance totale, comme le confirme la figure 5.19. On retrouve donc les conclusions de l'analyse spectrale du chapitre 4 et les résultats numériques obtenus sur l'illustration 1D de la section 4.3 et sur le problème de normalisation 1D de la section 5.2, et ce, même sur un problème 2D avec des modèles non-linéaires pour l'estimation d'un champ hétérogène.

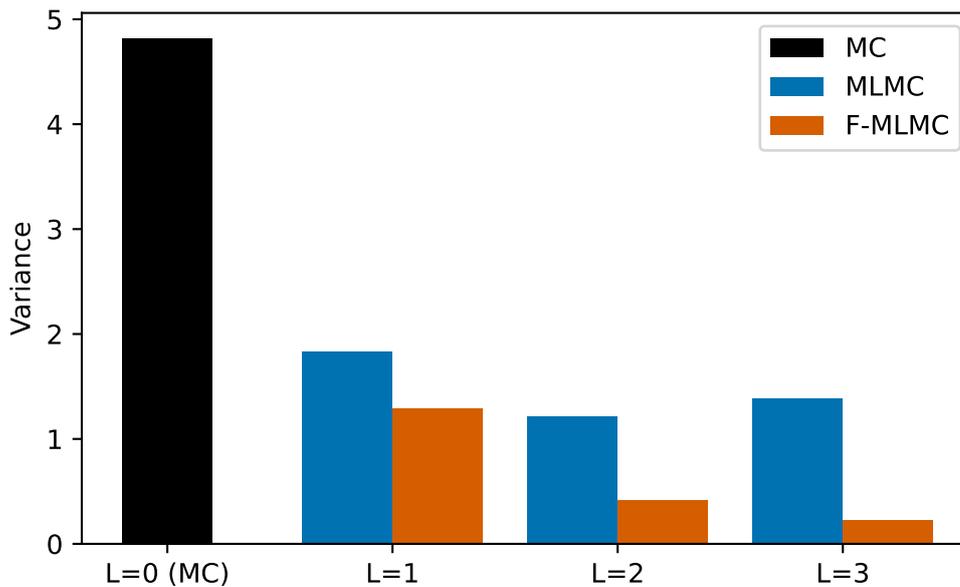


FIGURE 5.19 – Variance totale de l'estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC et F-MLMC de  $\mathbf{v}^{012}$ . Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.

On traite maintenant le problème de l'estimation de  $\mathbf{v}^{002}$ , basé sur le couple  $(\mathbf{d}_{11}, \mathbf{d}_{22})$  représentés sur la figure 5.12b, avec les paramètres  $\mu_\zeta = 0.02$  et  $D_\zeta = 0.04$ . Le vecteur référence  $\mathbf{v}^{002}$  calculé explicitement à partir de l'équation (5.4) est présenté sur la figure 5.20 en même temps que des exemples d'estimation MC, MLMC et F-MLMC à 2 et 4 niveaux.

Le champ  $\mathbf{v}^{002}$  contient bien plus de composantes fines échelles que  $\mathbf{v}^{012}$ , ce qui conduit à des estimations bien moins lisses. Les composantes parasites fines échelles introduites par les opérateurs de transfert dans l'estimation MLMC sont donc moins visibles, elles apparaissent à des échelles aussi fines que celles naturellement contenue dans  $\mathbf{v}^{002}$ . Néanmoins, par construction, les différents estimateurs estiment parfaitement  $\mathbf{v}^{002}$  en moyenne, comme montré sur la figure 5.21 avec une estimation de leur espérance comparée à la référence. Cela confirme une nouvelle fois que ces estimateurs sont non-biaisés et donc que leur MSE est uniquement constituée de leur variance. Les champs de variance sont affichés sur la figure 5.22.

De la même manière que pour les exemples d'estimation de la figure 5.20, les champs de variance des estimateurs MLMC contiennent des composantes fines échelles parasites, mais elles apparaissent aux mêmes échelles que celles contenues dans  $\mathbf{v}^{002}$ , les rendant difficiles à distinguer visuellement. Il est cependant possible d'observer que les champs de variance des estimateurs MLMC sont de norme plus

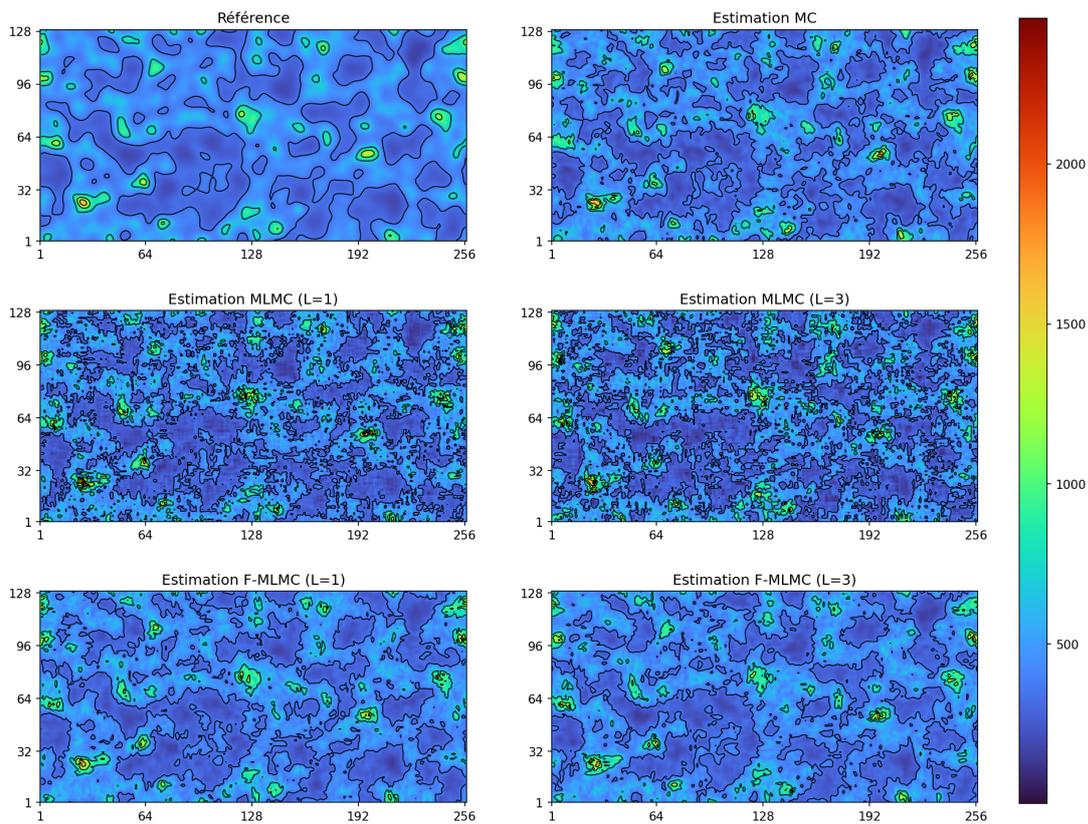


FIGURE 5.20 – Exemple d’estimation de  $v^{012}$  par un MLMC et un F-MLMC à 2 et 4 niveaux avec le budget total  $\eta = 100C_L$ , comparé à la référence  $v^{002}$ , en haut à gauche, et à une estimation par MC pour le même budget  $\eta$ , en haut à droite.

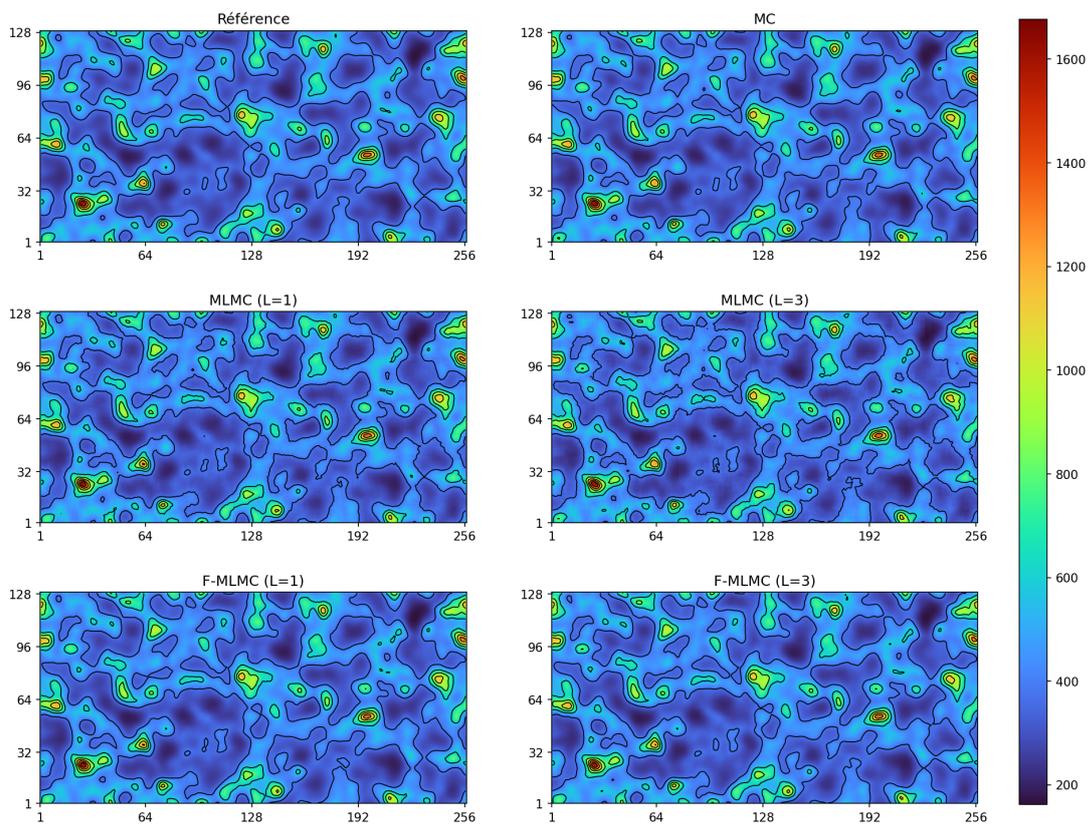


FIGURE 5.21 – Espérance des estimateurs MLMC et F-MLMC avec 2 et 4 niveaux, comparée à la référence  $v^{002}$  et à l’espérance de l’estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . L’espérance est estimée à partir de 100 estimateurs.

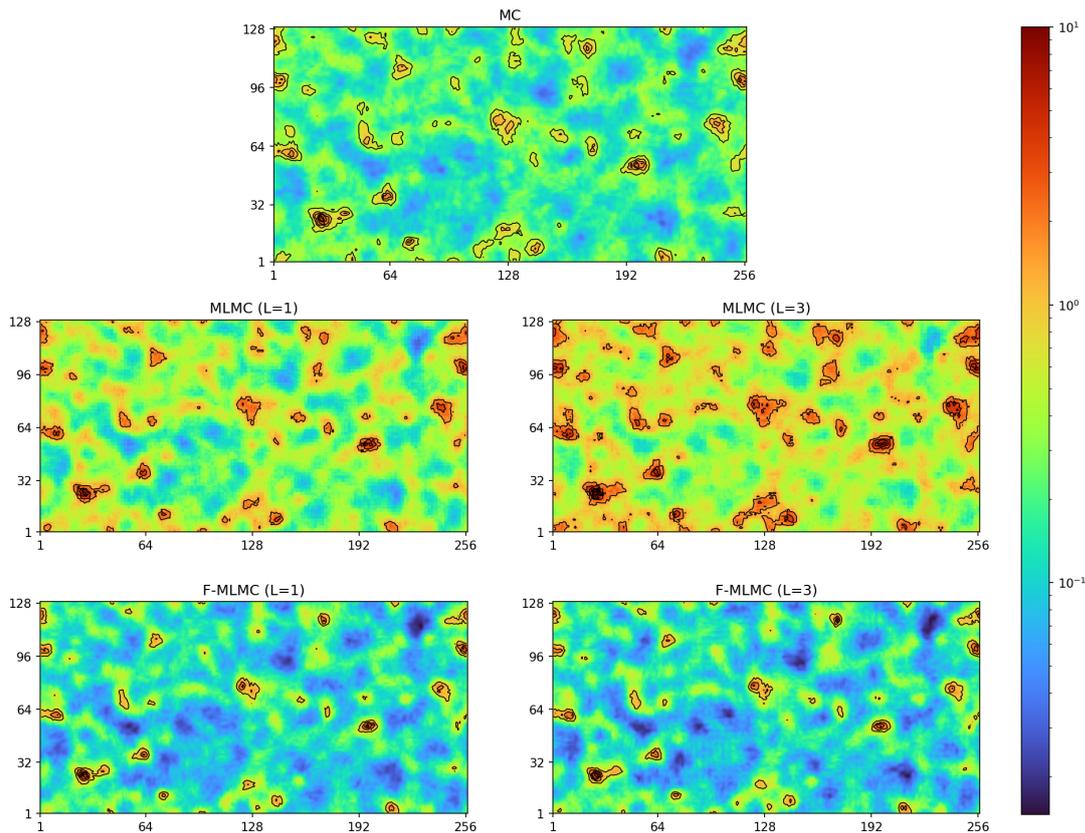


FIGURE 5.22 – Champs de variance des estimateurs MLMC et F-MLMC avec 2 et 4 niveaux de  $v^{002}$  comparés à celui de l’estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.

élevée que ceux du MC, tandis que la norme des champs de variance des estimateurs F-MLMC est la plus faible. Une meilleure visualisation de ces champs consiste donc à passer dans l’espace spectral de Hartley pour obtenir les champs de variance spectrale de la figure 5.23.

La première constatation est que la variance du MC est bien plus étalée sur les différentes composantes que celle de la figure 5.16. Les composantes moyennes et fines échelles étant plus importantes dans  $v^{002}$ , leur variance l’est également. L’utilisation de différents niveaux de fidélité par l’estimateur MLMC, et donc d’opérateurs de transfert associés, ajoute bien de l’erreur sur les composantes moyennes et fines échelles. Ces erreurs sont fortement réduites par l’ajout de filtres, donnant des profils de variance spectrales similaires entre le F-MLMC et le MC. On calcule les vecteurs de variance cumulée  $v^{cml}$  suivant l’équation (5.63) afin de mieux visualiser l’importance de ces erreurs par rapport à la variance totale et pour pouvoir comparer les erreurs sur les grandes échelles. Ces variances cumulées sont tracées sur la figure 5.24.

On remarque que l’erreur sur les composantes moyennes échelles contribuent, pour le MC, de manière plus importantes à la variance totale, donnant un profil de variance cumulée bien plus étalé. On constate également que les estimateurs MLMC atteignent des variances totales bien plus importantes que l’estimateur MC, ce qui est cohérent avec les observations faites sur la figure 5.22. L’ajout de niveaux de fidélité plus grossiers n’apporte pas assez de bénéfices comparé aux erreurs introduites par

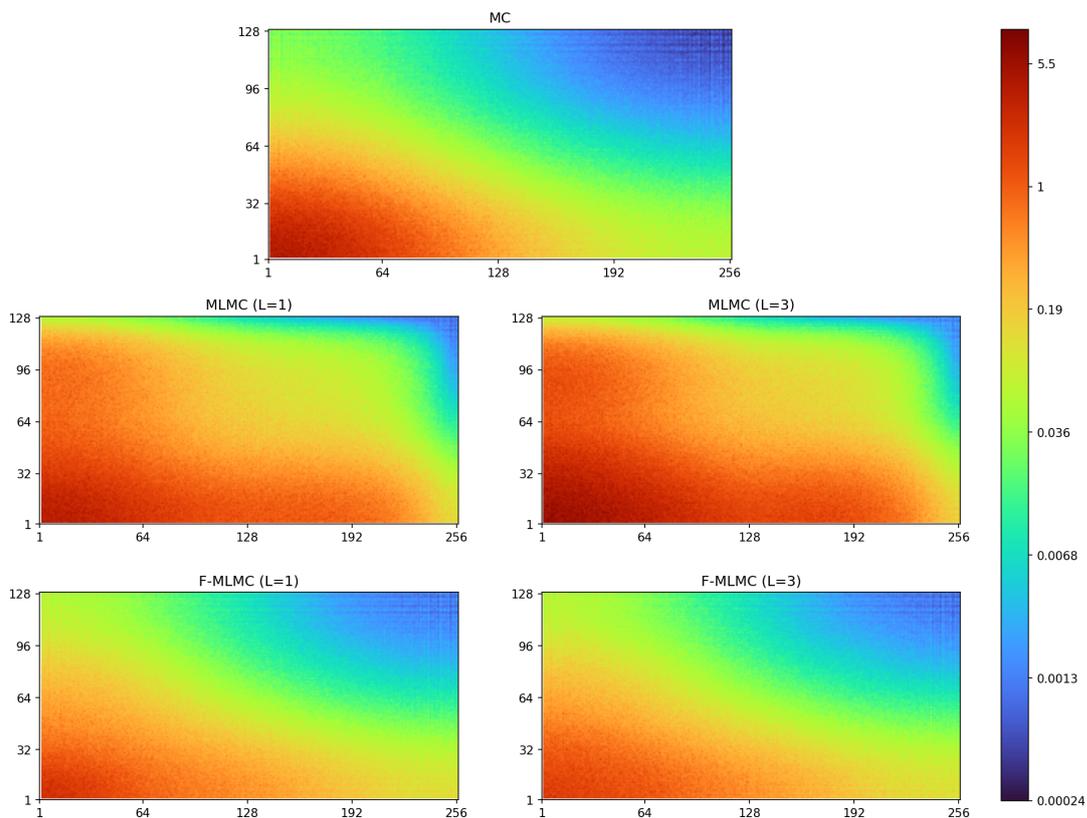
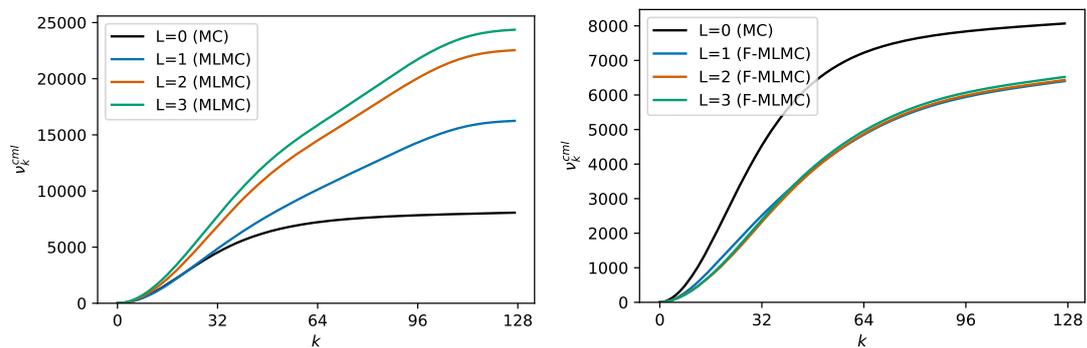


FIGURE 5.23 – Champs de variance spectrale des estimateurs MLMC et F-MLMC avec 2 et 4 niveaux de  $\mathbf{v}^{002}$  comparés à celui de l'estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.



(A) Variance cumulée des estimateurs MLMC.

(B) Variance cumulée des estimateurs F-MLMC.

FIGURE 5.24 – Variance cumulée  $\nu_k^{cml}$  des estimateurs MLMC (gauche) et F-MLMC (droite) avec  $L \in \{1, 2, 3\}$ , pour le problème d'estimation de  $\mathbf{v}^{002}$ , comparée à celle de l'estimateur MC. Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.

les opérateurs de transfert. Cela conduit à une détérioration de la variance sur la quasi-totalité du spectre, et donc à une hausse de la variance totale. Les courbes de variance cumulée des trois estimateurs F-MLMC sont quasiment identiques. L'utilisation de deux niveaux permet de construire un estimateur dont la variance est inférieure à celle du MC, mais l'ajout de niveaux supplémentaires ne permet pas plus d'améliorations. Cependant, les filtres permettent au moins d'atténuer fortement les erreurs apportées par les opérateurs de transfert, empêchant la variance totale d'augmenter comme c'est le cas pour le MLMC. Les variance totales obtenues par les différents estimateurs sont résumées dans la figure 5.25.

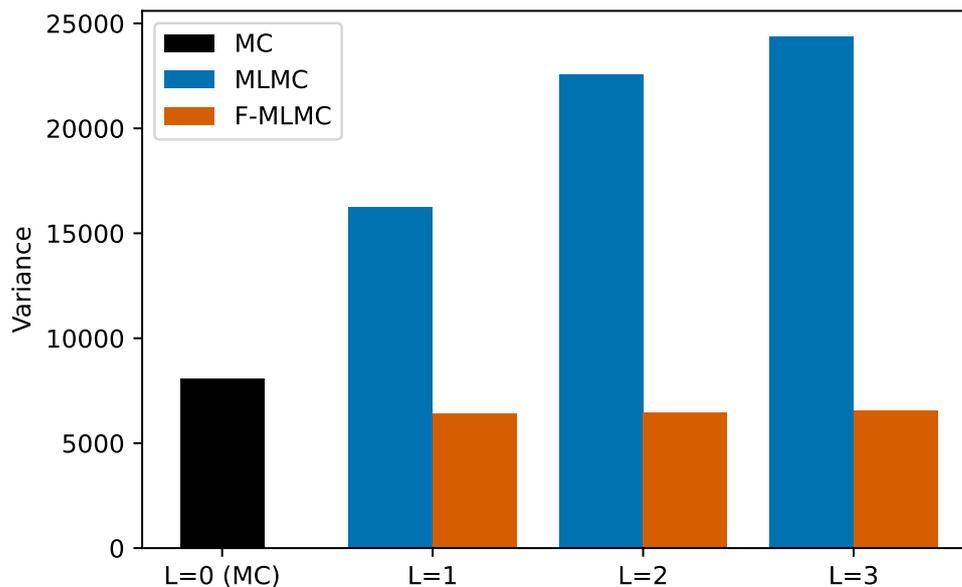


FIGURE 5.25 – Variance totale de l'estimateur MC ( $L = 0$ ) et de différents estimateurs MLMC et F-MLMC de  $\mathbf{v}^{002}$ . Le niveau le plus fin correspond à une grille de taille  $256 \times 128$  et le budget total est fixé à  $\eta = 100C_L$ . La variance est estimée à partir de 100 estimateurs.

## 5.4 Conclusion

Les chapitres 3 et 4 ont permis de montrer l'influence des opérateurs de transfert sur un estimateur MLMC et l'importance d'y ajouter des étapes de filtrage pour atténuer les erreurs introduites par ces opérateurs, et ce pour le cas simplifié de modèles linéaires, symétriques et circulants. Ce cas simplifié est loin d'être représentatif des applications que l'on peut trouver en géosciences, c'est pourquoi ce chapitre était dédié à des expériences numériques sur des problèmes plus complexes. En particulier, on s'est ici intéressé au problème de l'estimation des coefficients de normalisation d'un opérateur de covariance basé sur la diffusion en 1D et en 2D. Ce problème revient à estimer la diagonale d'une matrice de covariance et il est courant, en pratique, d'utiliser un estimateur MC pour approximer cette diagonale. Les expériences 1D ont d'abord permis d'étudier la variance des estimateurs MLMC et F-MLMC dans un cas où les modèles utilisés sont non-linéaires. L'analyse spectrale des chapitres 3 et 4 ne permet donc pas d'assurer que l'on y observera les mêmes comportements qu'avec des opérateurs linéaires, symétriques et circulants. Les résultats numériques sont, cependant, extrêmement similaires à ceux de l'illustration 1D de la section 4.3. On observe que les opérateurs de transfert introduisent des erreurs sur l'entièreté du spectre, mais

que l'utilisation de filtres permet d'atténuer fortement ces erreurs et donc d'obtenir un estimateur F-MLMC avec une meilleure variance totale. Les expériences 2D ont encore complexifiées les modèles utilisés en construisant les opérateurs de diffusion à partir d'un tenseur de diffusion hétérogène. Les coefficients de normalisation à estimer ne sont donc plus constants en espace. Cette complexification du problème n'a pas changé les conclusions, à savoir que l'ajout d'étapes de filtrage au sein des opérateurs de transfert permet une meilleure estimation de toutes les composantes d'échelle du champ discrétisé, et conduit donc à un estimateur avec une variance réduite comparée à celle du simple MLMC. Le choix des opérateurs de transfert pour un estimateur multi-niveaux basé sur des grilles de différentes résolution est donc une étape primordiale en vue de minimiser sa variance. Un tel choix est forcément dépendant du problème considéré, mais l'utilisation d'opérateurs incluant une étape de filtrage, ou d'interpolations d'ordre plus élevé, semble indispensable. En effet, on rappelle que l'ajout d'un filtre de Shapiro d'ordre 2 à l'opérateur de prolongation revient à interpoler linéairement (cf. équation (4.42)). On notera cependant certaines limites des expériences effectuées ici, l'utilisation d'échantillon pilote de grande taille pour estimer les quantités  $\mathcal{V}_\ell$ , nécessaire pour le calcul de l'allocation optimale des estimateurs, n'est pas souvent réalisable en pratique. De plus, le modèle des coûts de calcul utilisé ici n'est pas représentatif de simulateurs opérationnels. On pense notamment aux opérateurs de filtrage dont le coût d'application a été considéré négligeable. Finalement, le domaine spatial  $\mathcal{D}$  du problème de normalisation est bien plus complexe dans les modèles opérationnels [73], les conditions aux bords ne sont pas toujours périodiques et l'intérieur du domaine contient souvent des obstacles. Ces modifications demandent une réflexion plus poussée sur les opérateurs de transfert entre les grilles. Le test d'un estimateur F-MLMC sur un modèle opérationnel plus complexe permettrait de vérifier si ses avantages se maintiennent.

## Chapitre 6

# Conclusion et perspectives

L'utilisation d'estimateurs Monte Carlo multi-niveaux (MLMC) peut permettre, grâce à l'utilisation de modèles de différentes fidélités, d'atteindre une erreur quadratique moyenne bien plus faible qu'un estimateur Monte Carlo (MC) pour le même coût de calcul. Cette thèse s'est consacrée à étendre les applications du MLMC au domaine des géosciences. Spécifiquement, nous nous sommes concentrés sur l'étude du MLMC pour l'estimation de l'espérance d'un champ aléatoire discrétisé et sur l'utilisation d'une hiérarchie de grilles centrées sur les cellules pour définir les modèles de différentes fidélités. Les représentations discrétisées des champs étudiés sur les différentes grilles ont des dimensions inconsistantes, ce qui requiert diverses adaptations de l'estimateur.

Le chapitre 2 a d'abord permis de rappeler le fonctionnement et les propriétés du MLMC sur un cas simple d'estimation de l'espérance d'une variable aléatoire. L'adaptation du MLMC pour l'estimation de l'espérance d'un champ discrétisé par différence finie a été discutée dans le chapitre 3. L'utilisation de grilles de différentes résolutions requiert l'introduction d'opérateurs de transfert afin de pouvoir construire une hiérarchie de modèles dont les dimensions de toutes les entrées et sorties sont consistantes. Une analyse spectrale des opérateurs de transfert utilisés permet de montrer que ceux-ci introduisent des erreurs. L'opérateur de restriction transforme les composantes fines échelles en composantes grandes échelles, tandis que l'opérateur de prolongation introduit des composantes fines échelles qui parasitent le champ obtenu. L'impact de ces opérateurs sur la variance du MLMC a été exprimée dans le cas spécifique de modèles linéaires, symétriques et circulants, ce qui permet de constater que les erreurs introduites par les opérateurs de transfert influencent négativement cette variance. Ces résultats théoriques ont été confirmés par l'utilisation du MLMC sur une illustration 1D. Le MLMC atteint une variance plus faible qu'un simple MC grâce à une meilleure estimation des composantes grandes échelles. On observe, cependant, des erreurs bien plus importantes sur les fines échelles, erreurs pouvant contribuer de manière significative à la variance totale.

En s'inspirant des méthodes multigrilles, le chapitre 4 ajoute une étape de filtrage aux opérateurs de transfert, aboutissant à la construction de l'estimateur filtered-MLMC (F-MLMC). D'après les résultats de l'analyse spectrale, filtrer les fines échelles avant la restriction et après la prolongation atténue fortement les erreurs introduites par ces opérateurs. Cette réduction de variance a été quantifiée pour l'utilisation d'un filtre de Shapiro d'ordre 2 et dans le cas spécifique de modèles linéaires, symétriques et circulants. La comparaison du MLMC et du F-MLMC sur l'illustration 1D permet effectivement de constater les améliorations apportées par les filtres. Le F-MLMC obtient une variance plus faible sur l'ensemble des composantes du champ discrétisé, et donc une variance totale plus faible que le MLMC. Même dans les cas extrêmes où le MLMC détériore la variance comparé à un simple MC, le F-MLMC parvient à être meilleur que le MC. Le choix des opérateurs de transfert utilisés a donc un

impact fort sur les performances d'un estimateur MLMC et, sans savoir précisément quels sont les opérateurs optimaux, l'inclusion d'une étape de filtrage est essentielle. Il est également possible d'interpréter l'ajout de filtre comme l'augmentation de l'ordre des opérateurs de transfert, ce qui semble indiquer que, de la même manière que pour les méthodes multigrilles [33, 17], les opérateurs de transfert doivent être d'ordre suffisamment élevé. Les tests de plusieurs filtres et de plusieurs opérateurs de transfert ont permis de constater des gains de variance relativement similaires sur l'illustration 1D. Certains opérateurs d'ordre supérieur, tel que le filtre de Shapiro d'ordre 4 ou l'interpolation cubique, peuvent permettre de réduire un peu plus la variance de l'estimateur et l'utilisation d'un estimateur weighted MLMC (WMLMC) peut aider à définir les opérateurs de transfert à utiliser en donnant le coefficient multiplicatif optimal à appliquer devant ces opérateurs.

Finalement, le chapitre 5 applique les résultats précédents au problème de l'estimation des coefficients de normalisation d'un opérateur de covariance basé sur la diffusion. Malgré l'utilisation de modèles non-linéaires et le fait que le champ à estimer ne soit plus constant, l'estimateur F-MLMC obtient une meilleure estimation que le MLMC pour toutes les composantes spectrales et atteint donc toujours une variance totale plus faible. Les conclusions tirées de l'analyse spectrale concernant l'influence des opérateurs de transfert sur le MLMC, obtenus dans le cas spécifique de modèles linéaires, symétriques et circulants, semblent pouvoir se généraliser à des cas plus complexes et plus proches des applications rencontrées en géosciences.

Nous avons donc pu illustrer, dans cette thèse, l'importance du choix des opérateurs de transfert pour un estimateur MLMC de l'espérance d'un champ discrétisé lorsque les niveaux de fidélité se basent sur des grilles centrées sur les cellules. L'analyse spectrale effectuée a permis de mieux comprendre l'influence de ces opérateurs de transfert sur la variance de l'estimateur et d'en déduire que, de manière similaire aux méthodes multigrilles, l'ajout d'étapes de filtrage est primordiale. Les expériences numériques ont confirmé les améliorations apportées par l'estimateur F-MLMC par rapport au MLMC.

Ces travaux ouvrent plusieurs perspectives pour de futures recherches. En effet, si l'importance du choix des opérateurs de transfert a été démontrée, la problématique naturelle qui en découle porte alors sur le choix optimal de ces opérateurs. Un opérateur optimal serait un opérateur qui minimise la variance de l'estimateur MLMC (ou F-MLMC). L'utilisation du WMLMC dans la section 4.4.3 est un premier pas dans cette direction, mais d'autres estimateurs pourraient permettre de développer cette idée. On pense notamment à l'estimateur MLBLUE [63, 62] qui, par construction, combine les différents modèles de telle sorte que la variance soit minimale. Des extensions sur le MLBLUE pour l'estimation de quantités multi-variées [18, section 4.4] associent un poids scalaire à chaque composante de la sortie d'un modèle. Après un passage dans l'espace de Hartley, cela revient donc à associer un poids à chacune des composantes spectrales de la sortie d'un modèle. Un tel procédé pourrait permettre de trouver le filtre optimal à appliquer à la sortie d'un modèle ce qui, en plus de réduire la variance de l'estimateur, pourrait donner une meilleure compréhension des filtres à utiliser et des échelles que l'on souhaite conserver. Une autre piste pour aider au choix des opérateurs de transfert serait d'étudier la relation entre ordre de ces opérateurs et variance de l'estimateur. Cela pourrait permettre de trouver l'ordre des opérateurs qui minimise la variance de l'estimateur, de la même manière que [33] donne l'ordre des opérateurs de transfert nécessaire pour la convergence des méthodes multigrilles. L'ordre nécessaire peut alors être atteint par des applications successives de filtres de Shapiro et c'est, par exemple, ce qui a été fait dans [17], pour les méthodes multigrilles adaptées à l'assimilation de données dans un contexte géosciences. Finalement, et

de manière beaucoup plus simple, des tests plus complets de différents opérateurs sur différents problèmes aideraient à mieux comprendre quelles caractéristiques on souhaite avoir sur les opérateurs de transfert.

Outre le choix des opérateurs, différents points sur l'estimateur F-MLMC présenté dans cette thèse n'ont pas été traités. Premièrement, les avantages du F-MLMC ont été démontrés dans un cas encore relativement éloigné de nombreuses considérations pratiques. Par exemple, les coûts de calcul apportés par l'application des filtres ont été supposés négligables comparés aux coûts de l'application des modèles. Si, pour une certaine application, ce n'est pas le cas, prendre en compte le coût des filtres réduirait les performances du F-MLMC comparé à un simple MLMC. Le modèle de coût utilisé ici est également simplifié et le problème de l'allocation optimale sous contrainte de budget total fixé pourrait être amélioré [2]. De plus, le problème de l'allocation optimale repose sur une estimation de plusieurs termes de variances (ou de covariances pour le WMLMC) qui ont été estimés à l'aide d'un échantillon pilote de grande taille. Un tel procédé n'est pas souvent réaliste pour des applications pratiques, posant alors la question de l'influence de la qualité de ces estimations sur les performances de l'estimateur. Deuxièmement, seul le cas des grilles régulières a été considéré. L'utilisation de grilles non-uniformes demande une réflexion plus poussée sur les opérateurs de transfert et de filtrage à utiliser. Le traitement de grilles non-uniformes permettrait d'utiliser le F-MLMC dans bien plus d'applications en géosciences. La complexification du domaine d'étude peut aussi nécessiter des changements dans les opérateurs. On pense notamment à l'utilisation de conditions aux bords non-périodiques ou l'ajout d'obstacles au sein du domaine, par exemple des continents pour un problème d'océanographie. Dernièrement, un problème potentiellement important des estimateurs MLMC est que l'on ne peut pas garantir le caractère positif (ou négatif) de la quantité estimée [34]. Les coefficients de normalisation estimés dans l'application du chapitre 5 sont tous positifs, cependant, il est possible d'obtenir des estimations négatives de ces coefficients par MLMC. D'autres applications en géosciences concernent l'estimation d'une matrice de covariance, comme la matrice de covariance d'erreur d'ébauche  $\mathbf{B}$  en assimilation de données, et l'utilisation du MLMC pour estimer cette matrice ne permettrait pas de garantir son caractère semi-défini positif. Adresser ce problème pourrait encore étendre un peu plus les domaines d'applications du MLMC, et plusieurs travaux à ce sujet existent déjà [48, 49].

Une dernière perspective qui nous semble intéressante, particulièrement pour le problème de normalisation du chapitre 5, est la combinaison des méthodes MLMC et multigrilles. En effet, le problème de normalisation repose sur l'échantillonnage du vecteur aléatoire  $\mathbf{AX}$ , ce qui revient à résoudre de nombreux systèmes linéaires de la forme  $\mathbf{Mu} = \mathbf{b}$  où  $\mathbf{M}$  est symétrique définie positive. Dans nos expériences, la résolution de ces systèmes repose sur la décomposition de Cholesky qui n'est pas forcément réalisable en très grande dimension. On lui préférera alors l'utilisation de méthodes itératives [70]. Les méthodes multigrilles sont bien adaptées à un tel problème, d'autant plus que la hiérarchie de grilles et de modèles serait déjà construite pour l'estimateur MLMC. La possibilité de cumuler les gains apportés par les méthodes multigrilles pour la résolution de systèmes linéaires et les gains du MLMC (ou F-MLMC) pour l'estimation des coefficients de normalisation pourrait conduire à une amélioration considérable de la solution obtenue pour un budget fixé. De plus, l'utilisation des deux méthodes pour la résolution d'un même problème semble être un terrain intéressant pour étudier les similarités et possibles liens entre celles-ci, notamment au niveau des opérateurs de transfert et de filtrage utilisés.



## Annexe A

# Propriétés de la base de Hartley

### A.1 Orthogonalité de la matrice de Hartley

On considère ici une matrice de Hartley  $\mathbf{H}$  de taille  $n$ ,

$$(\mathbf{H})_{j,k} := \frac{1}{\sqrt{n}} (\cos \alpha_{jk} + \sin \alpha_{jk}), \quad \alpha_{jk} := \frac{(2j+1)k\pi}{n}, \quad \forall j, k = 0, \dots, n-1. \quad (\text{A.1})$$

On a, pour  $i, j = 0, \dots, n-1$ ,

$$(\mathbf{H}\mathbf{H}^T)_{i,j} = \frac{1}{n} \sum_{k=0}^{n-1} (\cos \alpha_{ik} + \sin \alpha_{ik})(\cos \alpha_{jk} + \sin \alpha_{jk}) \quad (\text{A.2})$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} [\cos(\alpha_{ik} - \alpha_{jk}) + \sin(\alpha_{ik} + \alpha_{jk})] \quad (\text{A.3})$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} \cos \frac{2(i-j)k\pi}{n} + \frac{1}{n} \sum_{k=0}^{n-1} \sin \frac{2(i+j+1)k\pi}{n}. \quad (\text{A.4})$$

À partir [28, 1.342 (1&2)], on déduit, pour  $i, j = 0, \dots, n-1$ ,

$$\sum_{k=0}^{n-1} \cos \frac{2(i-j)k\pi}{n} = n\delta_{ij}, \quad \sum_{k=0}^{n-1} \sin \frac{2(i+j+1)k\pi}{n} = 0, \quad (\text{A.5})$$

où  $\delta_{ij}$  est la notation du symbole delta de Kronecker, ce qui prouve bien  $\mathbf{H}\mathbf{H}^T = \mathbf{I}_n$ . De la même manière,

$$(\mathbf{H}^T\mathbf{H})_{i,j} = \frac{1}{n} \sum_{k=0}^{n-1} \cos \frac{(2k+1)(i-j)\pi}{n} + \frac{1}{n} \sum_{k=0}^{n-1} \sin \frac{(2k+1)(i+j)\pi}{n}, \quad (\text{A.6})$$

et on obtient  $\mathbf{H}^T\mathbf{H} = \mathbf{I}_n$  à l'aide de [28, 1.342 (3&4)].

### A.2 Les matrices symétriques circulantes sont diagonalisables dans la base de Hartley

Dans cette annexe, on prouve le théorème 3 ci-dessous. Ce théorème affirme que les matrices symétriques circulantes sont diagonalisables dans la base de Hartley centrée sur les cellules définie par l'équation (A.1). Pour cela, on commence par rappeler ou prouver ces résultats sur la base de Fourier centrée sur les nœuds  $\check{\mathbf{F}} \in \mathbb{C}^{n \times n}$

et sur la base de Hartley centrée sur les nœuds  $\check{\mathbf{H}}^\pm \in \mathbb{R}^{n \times n}$ , définies par

$$\check{\mathbf{F}} := \check{\mathbf{H}}_c + i\check{\mathbf{H}}_s, \quad \check{\mathbf{H}}^\pm := \check{\mathbf{H}}_c \pm \check{\mathbf{H}}_s, \quad (\text{A.7})$$

$$(\check{\mathbf{H}}_c)_{j,k} := \frac{1}{\sqrt{n}} \cos \frac{2jk\pi}{n}, \quad (\check{\mathbf{H}}_s)_{j,k} := \frac{1}{\sqrt{n}} \sin \frac{2jk\pi}{n}, \quad (\text{A.8})$$

où  $i \in \mathbb{C}$  représente le nombre complexe unitaire  $i^2 = -1$ . Il est clair que  $\check{\mathbf{H}}_c$  et  $\check{\mathbf{H}}_s$  sont des matrices symétriques réelles, et donc  $\check{\mathbf{H}}^\pm$  l'est également.  $\check{\mathbf{F}}$ , cependant, est une matrice symétrique complexe et non-Hermitienne. De plus,  $\check{\mathbf{F}}$  est unitaire (cf. [16]), i.e.,  $\check{\mathbf{F}}^* \check{\mathbf{F}} = \check{\mathbf{F}} \check{\mathbf{F}}^* = \mathbf{I}_n$ , où  $\check{\mathbf{F}}^* = \check{\mathbf{H}}_c - i\check{\mathbf{H}}_s$  est la transposée conjuguée de  $\check{\mathbf{F}}$ .

**Lemme 1.**  $\check{\mathbf{H}}_c$  et  $\check{\mathbf{H}}_s$  sont telles que

$$\check{\mathbf{H}}_c^2 + \check{\mathbf{H}}_s^2 = \mathbf{I}_n, \quad \check{\mathbf{H}}_c \check{\mathbf{H}}_s = \check{\mathbf{H}}_s \check{\mathbf{H}}_c = \mathbf{0}_n, \quad \check{\mathbf{H}}_c \mathbf{1}_n = \sqrt{n} \mathbf{e}_1, \quad \check{\mathbf{H}}_s \mathbf{1}_n = \mathbf{0}_n, \quad (\text{A.9})$$

où  $\mathbf{0}_n := (0, \dots, 0)^T \in \mathbb{R}^n$ ,  $\mathbf{1}_n := (1, \dots, 1)^T \in \mathbb{R}^n$ , et  $\mathbf{e}_1$  est la première colonne de  $\mathbf{I}_n$ .

**Preuve.** Les deux premières identités sont tirées de [5, Lemme 1], tandis que les deux autres viennent de [28, 1.342 (1&2)].

**Corollaire 1.**  $\check{\mathbf{H}}^\pm \mathbf{1}_n = \sqrt{n} \mathbf{e}_1$ .

**Corollaire 2.**  $\check{\mathbf{H}}^\pm$  est orthogonale, i.e.,  $(\check{\mathbf{H}}^\pm)^2 = \mathbf{I}_n$ . De plus,  $\check{\mathbf{H}}^+ \check{\mathbf{H}}^- = \check{\mathbf{H}}^- \check{\mathbf{H}}^+ = \check{\mathbf{F}}^2$ .

**Preuve.** Avec les définitions de l'équation (A.7) et le lemme 1, on a

$$\check{\mathbf{H}}^\pm \mathbf{1}_n = \check{\mathbf{H}}_c \mathbf{1}_n \pm \check{\mathbf{H}}_s \mathbf{1}_n = \check{\mathbf{H}}_c \mathbf{1}_n = \sqrt{n} \mathbf{e}_1, \quad (\text{A.10})$$

$$(\check{\mathbf{H}}^\pm)^2 = (\check{\mathbf{H}}_c^2 + \check{\mathbf{H}}_s^2) \pm (\check{\mathbf{H}}_c \check{\mathbf{H}}_s + \check{\mathbf{H}}_s \check{\mathbf{H}}_c) = \check{\mathbf{H}}_c^2 + \check{\mathbf{H}}_s^2 = \mathbf{I}_n, \quad (\text{A.11})$$

$$\check{\mathbf{F}}^2 = \check{\mathbf{H}}_c^2 - \check{\mathbf{H}}_s^2 + i(\check{\mathbf{H}}_c \check{\mathbf{H}}_s + \check{\mathbf{H}}_s \check{\mathbf{H}}_c) = \check{\mathbf{H}}_c^2 - \check{\mathbf{H}}_s^2, \quad (\text{A.12})$$

$$\check{\mathbf{H}}^- \check{\mathbf{H}}^+ = \check{\mathbf{H}}_c^2 - \check{\mathbf{H}}_s^2 + (\check{\mathbf{H}}_c \check{\mathbf{H}}_s - \check{\mathbf{H}}_s \check{\mathbf{H}}_c) = \check{\mathbf{F}}^2 = \check{\mathbf{H}}^+ \check{\mathbf{H}}^-. \quad (\text{A.13})$$

**Lemme 2.** Soit  $\mathbf{A} = \text{Circ}(\mathbf{a}) \in \mathbb{R}^{n \times n}$  une matrice symétrique circulante dont la première colonne est  $\mathbf{a} = (a_k)_{k=0}^{n-1} \in \mathbb{R}^n$ , avec  $a_{n-i} = a_i$ , pour  $i = 1, \dots, n-1$ . On a  $\check{\mathbf{H}}_c \mathbf{A} \check{\mathbf{H}}_s + \check{\mathbf{H}}_s \mathbf{A} \check{\mathbf{H}}_c = \check{\mathbf{H}}_c \mathbf{A} \check{\mathbf{H}}_s - \check{\mathbf{H}}_s \mathbf{A} \check{\mathbf{H}}_c = \mathbf{0}_n$ .

**Preuve.** Prouvé dans les étapes intermédiaires de la preuve du [5, Théorème 1].

**Théorème 2.** Soit  $\mathbf{A} = \text{Circ}(\mathbf{a}) \in \mathbb{R}^{n \times n}$  une matrice symétrique circulante dont la première colonne est  $\mathbf{a} = (a_k)_{k=0}^{n-1} \in \mathbb{R}^n$ , avec  $a_{n-i} = a_i$ , pour  $i = 1, \dots, n-1$ . On a  $\check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^+ = \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^- = \mathbf{\Lambda}$ , où  $\mathbf{\Lambda} = \sqrt{n} \text{Diag}(\check{\mathbf{H}}^+ \mathbf{a}) = \sqrt{n} \text{Diag}(\check{\mathbf{H}}^- \mathbf{a}) = \sqrt{n} \text{Diag}(\check{\mathbf{H}}_c \mathbf{a})$ .

**Preuve.** Le fait que  $\check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^+ = \sqrt{n} \text{Diag}(\check{\mathbf{H}}^+ \mathbf{a}) = \mathbf{\Lambda} = \check{\mathbf{H}}_c \mathbf{A} \check{\mathbf{H}}_c + \check{\mathbf{H}}_s \mathbf{A} \check{\mathbf{H}}_s$  provient du [5, Théorème 1], car les matrices symétriques circulantes appartiennent à la classe plus large des matrices considérées dans [5]. En partant de la définition de  $\check{\mathbf{H}}^-$  et du lemme 2, on a

$$\check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^- = (\check{\mathbf{H}}_c \mathbf{A} \check{\mathbf{H}}_c + \check{\mathbf{H}}_s \mathbf{A} \check{\mathbf{H}}_s) - (\check{\mathbf{H}}_c \mathbf{A} \check{\mathbf{H}}_s + \check{\mathbf{H}}_s \mathbf{A} \check{\mathbf{H}}_c) = \check{\mathbf{H}}_c \mathbf{A} \check{\mathbf{H}}_c + \check{\mathbf{H}}_s \mathbf{A} \check{\mathbf{H}}_s = \mathbf{\Lambda}. \quad (\text{A.14})$$

Finalement, à partir du lemme 1 et du corollaire 1,

$$\text{diag}(\mathbf{\Lambda}) = \mathbf{\Lambda} \mathbf{1}_n = \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^- \mathbf{1}_n = \check{\mathbf{H}}_c \mathbf{A} \check{\mathbf{H}}_c \mathbf{1}_n + \check{\mathbf{H}}_s \mathbf{A} \check{\mathbf{H}}_s \mathbf{1}_n = \sqrt{n} \check{\mathbf{H}}^- \mathbf{A} \mathbf{e}_1 = \sqrt{n} \check{\mathbf{H}}_c \mathbf{A} \mathbf{e}_1, \quad (\text{A.15})$$

ce qui conclut la preuve, car  $\mathbf{A} \mathbf{e}_1 = \mathbf{a}$ .

**Corollaire 3.** Soit  $\mathbf{A} = \text{Circ}(\mathbf{a}) \in \mathbb{R}^{n \times n}$  une matrice symétrique circulante et soit  $\Lambda$  définie comme dans le théorème 2. Alors  $\check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^- = \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^+ = \Lambda \check{\mathbf{F}}^2 = \check{\mathbf{F}}^2 \Lambda$ .

**Preuve.** À partir du théorème 2 et du corollaire 2, on a

$$\check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^- = \check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^+ \check{\mathbf{H}}^- = \check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^+ \check{\mathbf{F}}^2 = \Lambda \check{\mathbf{F}}^2, \quad (\text{A.16})$$

$$\check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^+ = \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^- \check{\mathbf{H}}^+ = \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^- \check{\mathbf{F}}^2 = \Lambda \check{\mathbf{F}}^2, \quad (\text{A.17})$$

$$(\check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^-)^T = (\Lambda \check{\mathbf{F}}^2)^T = \check{\mathbf{F}}^2 \Lambda = \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^+ = \Lambda \check{\mathbf{F}}^2, \quad (\text{A.18})$$

où la dernière ligne s'obtient grâce à la symétrie de  $\mathbf{A}$ ,  $\check{\mathbf{H}}^\pm$  et  $\check{\mathbf{F}}^2$

**Lemme 3.**  $\mathbf{H} = \check{\mathbf{H}}^+ \mathbf{C} + \check{\mathbf{H}}^- \mathbf{S}$ , où  $\mathbf{C} := \text{Diag}(\{\cos \frac{k\pi}{n}\}_{k=0}^{n-1})$  et  $\mathbf{S} := \text{Diag}(\{\sin \frac{k\pi}{n}\}_{k=0}^{n-1})$ .

**Preuve.** En utilisant les relations trigonométriques élémentaires, on a

$$\cos \frac{(2j+1)k\pi}{n} = \cos \frac{2jk\pi}{n} \cos \frac{k\pi}{n} - \sin \frac{2jk\pi}{n} \sin \frac{k\pi}{n}, \quad (\text{A.19})$$

$$\sin \frac{(2j+1)k\pi}{n} = \sin \frac{2jk\pi}{n} \cos \frac{k\pi}{n} - \cos \frac{2jk\pi}{n} \sin \frac{k\pi}{n}, \quad (\text{A.20})$$

et le lemme 3 s'ensuit.

**Lemme 4.**  $\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S} + \mathbf{S} \check{\mathbf{F}}^2 \mathbf{C} = \mathbf{0}_n$ , où  $\mathbf{C}$  et  $\mathbf{S}$  sont définies comme dans le lemme 3.

**Preuve.** À partir du [5, Lemme 1], on a que, pour  $j, k = 0, \dots, n-1$ ,  $(\check{\mathbf{F}}^2)_{j,k} = 1$  si  $j = k = 0$  ou  $j + k = n$ , et  $(\check{\mathbf{F}}^2)_{j,k} = 0$  sinon. De plus,  $(\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})_{j,k} = (\check{\mathbf{F}}^2)_{j,k} \cos \frac{j\pi}{n} \sin \frac{k\pi}{n}$ , et, par symétrie,  $\mathbf{S} \check{\mathbf{F}}^2 \mathbf{C} = (\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})^T$ . Maintenant, si  $j = k = 0$ ,  $(\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})_{j,k} = (\mathbf{S} \check{\mathbf{F}}^2 \mathbf{C})_{j,k} = 0$  trivialement. Sinon, si  $j \neq 0$ ,  $k \neq 0$ , et  $j + k \neq n$ ,  $(\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})_{j,k} = (\mathbf{S} \check{\mathbf{F}}^2 \mathbf{C})_{j,k} = 0$ . Finalement, si  $j + k = n$ , les relations trigonométriques élémentaires donnent

$$(\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})_{j,k} = \cos \frac{j\pi}{n} \sin \frac{(n-j)\pi}{n} = \cos \frac{j\pi}{n} \sin \left( \pi - \frac{j\pi}{n} \right) = \cos \frac{j\pi}{n} \sin \frac{j\pi}{n}, \quad (\text{A.21})$$

$$(\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})_{k,j} = \sin \frac{j\pi}{n} \cos \frac{(n-j)\pi}{n} = \sin \frac{j\pi}{n} \cos \left( \pi - \frac{j\pi}{n} \right) = -\cos \frac{j\pi}{n} \sin \frac{j\pi}{n}, \quad (\text{A.22})$$

et on a  $(\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S} + \mathbf{S} \check{\mathbf{F}}^2 \mathbf{C})_{j,k} = (\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})_{j,k} + (\mathbf{C} \check{\mathbf{F}}^2 \mathbf{S})_{k,j} = 0$ , ce qui conclut la preuve.

On peut maintenant poser et prouver le théorème 3.

**Théorème 3.** Soit  $\mathbf{A} = \text{Circ}(\mathbf{a}) \in \mathbb{R}^{n \times n}$  une matrice symétrique circulante et soit  $\Lambda$  définie comme dans le théorème 2. Alors  $\mathbf{H}^T \mathbf{A} \mathbf{H} = \Lambda$ .

**Preuve.** En partant de l'expression du lemme 3, et en appliquant le théorème 2, le corollaire 3 et le lemme 4, on obtient

$$\mathbf{H}^T \mathbf{A} \mathbf{H} = \mathbf{C} \check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^+ \mathbf{C} + \mathbf{S} \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^- \mathbf{S} + \mathbf{C} \check{\mathbf{H}}^+ \mathbf{A} \check{\mathbf{H}}^- \mathbf{S} + \mathbf{S} \check{\mathbf{H}}^- \mathbf{A} \check{\mathbf{H}}^+ \mathbf{C} \quad (\text{A.23})$$

$$= \mathbf{C} \Lambda \mathbf{C} + \mathbf{S} \Lambda \mathbf{S} + \mathbf{C} \Lambda \check{\mathbf{F}}^2 \mathbf{S} + \mathbf{S} \Lambda \check{\mathbf{F}}^2 \mathbf{C} = \Lambda (\mathbf{C}^2 + \mathbf{S}^2 + \mathbf{C} \check{\mathbf{F}}^2 \mathbf{S} + \mathbf{S} \check{\mathbf{F}}^2 \mathbf{C}), \quad (\text{A.24})$$

tel que  $\mathbf{H}^T \mathbf{A} \mathbf{H} = \Lambda (\mathbf{C}^2 + \mathbf{S}^2) = \Lambda$  car  $\mathbf{C}^2 + \mathbf{S}^2 = \mathbf{I}_n$  par trigonométrie élémentaire.



# Bibliographie

- [1] Mohamed Reda El AMRI et al. *Multilevel Surrogate-based Control Variates*. Juin 2023. arXiv : [2306.10800](https://arxiv.org/abs/2306.10800) [math, stat]. (Visité le 29/06/2023).
- [2] Niklas BAUMGARTEN, Sebastian KRUMSCHEID et Christian WIENERS. *A Fully Parallelized and Budgeted Multi-Level Monte Carlo Method and the Application to Acoustic Waves*. Juill. 2023. DOI : [10.48550/arXiv.2307.10767](https://doi.org/10.48550/arXiv.2307.10767). arXiv : [2307.10767](https://arxiv.org/abs/2307.10767) [cs, math]. (Visité le 18/06/2024).
- [3] Claudio BIERIG et Alexey CHERNOV. « Convergence Analysis of Multilevel Monte Carlo Variance Estimators and Application for Random Obstacle Problems ». In : *Numerische Mathematik* 130.4 (août 2015), p. 579-613. ISSN : 0945-3245. DOI : [10.1007/s00211-014-0676-3](https://doi.org/10.1007/s00211-014-0676-3). (Visité le 07/06/2022).
- [4] Claudio BIERIG et Alexey CHERNOV. « Estimation of Arbitrary Order Central Statistical Moments by the Multilevel Monte Carlo Method ». In : *Stochastics and Partial Differential Equations Analysis and Computations* 4.1 (mars 2016), p. 3-40. ISSN : 2194-041X. DOI : [10.1007/s40072-015-0063-9](https://doi.org/10.1007/s40072-015-0063-9). (Visité le 21/06/2023).
- [5] Dario BINI et Paola FAVATI. « On a Matrix Algebra Related to the Discrete Hartley Transform ». In : *SIAM Journal on Matrix Analysis and Applications* 14.2 (avr. 1993), p. 500-507. ISSN : 0895-4798. DOI : [10.1137/0614035](https://doi.org/10.1137/0614035). (Visité le 24/01/2023).
- [6] R. N. BRACEWELL. « Discrete Hartley Transform ». In : *JOSA* 73.12 (déc. 1983), p. 1832-1835. DOI : [10.1364/JOSA.73.001832](https://doi.org/10.1364/JOSA.73.001832). (Visité le 01/08/2024).
- [7] William L. BRIGGS, Van Emden HENSON et Steve F. MCCORMICK. *A Multigrid Tutorial, Second Edition*. Second. Society for Industrial and Applied Mathematics, jan. 2000. ISBN : 978-0-89871-462-3 978-0-89871-950-5. DOI : [10.1137/1.9780898719505](https://doi.org/10.1137/1.9780898719505). (Visité le 12/07/2022).
- [8] Richard H. BYRD, Mary E. HRIBAR et Jorge NOCEDAL. « An Interior Point Algorithm for Large-Scale Nonlinear Programming ». In : *SIAM Journal on Optimization* 9.4 (jan. 1999), p. 877-900. ISSN : 1052-6234. DOI : [10.1137/S1052623497325107](https://doi.org/10.1137/S1052623497325107). (Visité le 04/08/2024).
- [9] Russel E. CAFLISCH. « Monte Carlo and Quasi-Monte Carlo Methods ». In : *Acta Numerica* 7 (jan. 1998), p. 1-49. ISSN : 1474-0508, 0962-4929. DOI : [10.1017/S0962492900002804](https://doi.org/10.1017/S0962492900002804). (Visité le 21/06/2023).
- [10] Alberto CARRASSI et al. « Data Assimilation in the Geosciences : An Overview of Methods, Issues, and Perspectives ». In : *WIREs Climate Change* 9.5 (2018), e535. ISSN : 1757-7799. DOI : [10.1002/wcc.535](https://doi.org/10.1002/wcc.535). (Visité le 04/01/2023).
- [11] Alexey CHERNOV et al. « Multilevel Ensemble Kalman Filtering for Spatio-Temporal Processes ». In : *arXiv.org* (oct. 2017). (Visité le 21/06/2023).
- [12] Andrew K. CLIFFE et al. « Multilevel Monte Carlo Methods and Applications to Elliptic PDEs with Random Coefficients ». In : *Computing and Visualization in Science* 14.1 (août 2011), p. 3. ISSN : 1433-0369. DOI : [10.1007/s00791-011-0160-x](https://doi.org/10.1007/s00791-011-0160-x). (Visité le 04/01/2023).

- [13] Matteo CROCI, Michael B. GILES et Patrick E. FARRELL. « Multilevel Quasi Monte Carlo Methods for Elliptic PDEs with Random Field Coefficients via Fast White Noise Sampling ». In : *SIAM Journal on Scientific Computing* 43.4 (jan. 2021), A2840-A2868. ISSN : 1064-8275. DOI : [10.1137/20M1329044](https://doi.org/10.1137/20M1329044). (Visité le 20/02/2024).
- [14] Matteo CROCI, Karen E. WILLCOX et Stephen J. WRIGHT. « Multi-Output Multilevel Best Linear Unbiased Estimators via Semidefinite Programming ». In : (jan. 2023). DOI : [10.48550/arXiv.2301.07831](https://doi.org/10.48550/arXiv.2301.07831). (Visité le 31/01/2023).
- [15] Matteo CROCI et al. « Efficient White Noise Sampling and Coupling for Multilevel Monte Carlo with Nonnested Meshes ». In : *SIAM/ASA Journal on Uncertainty Quantification* 6.4 (jan. 2018), p. 1630-1655. DOI : [10.1137/18M1175239](https://doi.org/10.1137/18M1175239). (Visité le 14/02/2023).
- [16] Philip J. DAVIS. *Circulant Matrices : Second Edition*. 2nd. American Mathematical Society, 2012. ISBN : 978-0-8218-9165-0.
- [17] Laurent DEBREU et al. « Multigrid Solvers and Multigrid Preconditioners for the Solution of Variational Data Assimilation Problems ». In : *Quarterly Journal of the Royal Meteorological Society* 142.694 (2016), p. 515-528. ISSN : 1477-870X. DOI : [10.1002/qj.2676](https://doi.org/10.1002/qj.2676). (Visité le 07/04/2023).
- [18] Mayeul DESTOUCHES, Paul MYCEK et Selime GÜROL. *Multivariate Extensions of the Multilevel Best Linear Unbiased Estimator for Ensemble-Variational Data Assimilation*. Juin 2023. arXiv : [2306.07017](https://arxiv.org/abs/2306.07017) [[physics](#), [stat](#)]. (Visité le 29/06/2023).
- [19] Y. EFENDIEV, O. ILIEV et C. KRONSBELN. « Multilevel Monte Carlo Methods Using Ensemble Level Mixed MsFEM for Two-Phase Flow and Transport Simulations ». In : *Computational Geosciences* 17.5 (oct. 2013), p. 833-850. ISSN : 1573-1499. DOI : [10.1007/s10596-013-9358-y](https://doi.org/10.1007/s10596-013-9358-y). (Visité le 23/03/2024).
- [20] F. FALISSARD. « Genuinely Multi-Dimensional Explicit and Implicit Generalized Shapiro Filters for Weather Forecasting, Computational Fluid Dynamics and Aeroacoustics ». In : *Journal of Computational Physics* 253 (nov. 2013), p. 344-367. ISSN : 0021-9991. DOI : [10.1016/j.jcp.2013.07.001](https://doi.org/10.1016/j.jcp.2013.07.001). (Visité le 04/01/2023).
- [21] F. FALISSARD. « Uneven-Order Decentered Shapiro Filters for Boundary Filtering ». In : *Journal of Computational Physics* 292 (juill. 2015), p. 168-175. ISSN : 0021-9991. DOI : [10.1016/j.jcp.2015.03.003](https://doi.org/10.1016/j.jcp.2015.03.003). (Visité le 04/01/2023).
- [22] Gianluca GERACI, Michael S. ELDRED et Gianluca IACCARINO. « A Multifidelity Multilevel Monte Carlo Method for Uncertainty Propagation in Aerospace Applications ». In : *19th AIAA Non-Deterministic Approaches Conference*. American Institute of Aeronautics and Astronautics, 2017. DOI : [10.2514/6.2017-1951](https://doi.org/10.2514/6.2017-1951). (Visité le 30/05/2023).
- [23] Michael B. GILES. « Multilevel Monte Carlo Methods ». In : *Acta Numerica* 24 (mai 2015), p. 259-328. ISSN : 0962-4929, 1474-0508. DOI : [10.1017/S096249291500001X](https://doi.org/10.1017/S096249291500001X). (Visité le 07/06/2022).
- [24] Michael B. GILES. « Multilevel Monte Carlo Path Simulation ». In : *Operations Research* 56.3 (juin 2008), p. 607-617. ISSN : 0030-364X. DOI : [10.1287/opre.1070.0496](https://doi.org/10.1287/opre.1070.0496). (Visité le 04/01/2023).
- [25] Michael B. GILES et Lukasz SZPRUCH. « Multilevel Monte Carlo Methods for Applications in Finance ». In : *Recent Developments in Computational Finance*. T. Volume 14. Interdisciplinary Mathematical Sciences. WORLD SCIENTIFIC, août 2012, p. 3-47. ISBN : 978-981-4436-42-7. DOI : [10.1142/9789814436434\\_0001](https://doi.org/10.1142/9789814436434_0001). (Visité le 23/03/2024).

- [26] Michael B. GILES et Benjamin WATERHOUSE. « Multilevel Quasi-Monte Carlo Path Simulation ». In : *Advanced Financial Modelling*. Sous la dir. d'Hansjörg ALBRECHER, Wolfgang J. RUNGALDIER et Walter SCHACHERMAYER. Walter de Gruyter, oct. 2009, p. 165-182. ISBN : 978-3-11-021313-3. DOI : [10.1515/9783110213140.165](https://doi.org/10.1515/9783110213140.165). (Visité le 23/03/2024).
- [27] Takashi GODA et al. « Decision-Theoretic Sensitivity Analysis for Reservoir Development under Uncertainty Using Multilevel Quasi-Monte Carlo Methods ». In : *Computational Geosciences* 22.4 (août 2018), p. 1009-1020. ISSN : 1573-1499. DOI : [10.1007/s10596-018-9735-7](https://doi.org/10.1007/s10596-018-9735-7). (Visité le 23/03/2024).
- [28] I. S. GRADSHTEYN et I. M. RYZHIK. *Table of Integrals, Series, and Products*. Academic Press, mai 2014. ISBN : 978-1-4832-6564-3.
- [29] Peter GUTTORP et Tilmann GNEITING. « Studies in the History of Probability and Statistics XLIX On the Matérn Correlation Family ». In : *Biometrika* 93.4 (déc. 2006), p. 989-995. ISSN : 0006-3444. DOI : [10.1093/biomet/93.4.989](https://doi.org/10.1093/biomet/93.4.989). (Visité le 23/02/2024).
- [30] Wolfgang HACKBUSCH. *Multi-Grid Methods and Applications*. T. 4. Springer Series in Computational Mathematics. Berlin, Heidelberg : Springer, 1985. ISBN : 978-3-642-05722-9 978-3-662-02427-0. DOI : [10.1007/978-3-662-02427-0](https://doi.org/10.1007/978-3-662-02427-0). (Visité le 23/03/2024).
- [31] S. HEINRICH. « Monte Carlo Complexity of Global Solution of Integral Equations ». In : *Journal of Complexity* 14.2 (juin 1998), p. 151-175. ISSN : 0885-064X. DOI : [10.1006/jcom.1998.0471](https://doi.org/10.1006/jcom.1998.0471). (Visité le 28/06/2023).
- [32] Stefan HEINRICH. « Multilevel Monte Carlo Methods ». In : *Large-Scale Scientific Computing*. Sous la dir. de Svetozar MARGENOV, Jerzy WAŚNIEWSKI et Plamen YALAMOV. Lecture Notes in Computer Science. Berlin, Heidelberg : Springer, 2001, p. 58-67. ISBN : 978-3-540-45346-8. DOI : [10.1007/3-540-45346-6\\_5](https://doi.org/10.1007/3-540-45346-6_5).
- [33] P. W. HEMKER. « On the Order of Prolongations and Restrictions in Multigrid Procedures ». In : *Journal of Computational and Applied Mathematics* 32.3 (déc. 1990), p. 423-429. ISSN : 0377-0427. DOI : [10.1016/0377-0427\(90\)90047-4](https://doi.org/10.1016/0377-0427(90)90047-4). (Visité le 02/05/2024).
- [34] Håkon HOEL, Kody J. H. LAW et Raul TEMPONE. « Multilevel Ensemble Kalman Filtering ». In : *SIAM Journal on Numerical Analysis* 54.3 (jan. 2016), p. 1813-1839. ISSN : 0036-1429. DOI : [10.1137/15M100955X](https://doi.org/10.1137/15M100955X). (Visité le 04/01/2023).
- [35] Anastasia ISTRATUCA et Aretha TECKENTRUP. *Smoothed Circulant Embedding with Applications to Multilevel Monte Carlo Methods for PDEs with Random Coefficients*. Juin 2023. arXiv : [2306.13493](https://arxiv.org/abs/2306.13493) [cs, math]. (Visité le 30/10/2023).
- [36] M KHALIL et P WESSELING. « Vertex-Centered and Cell-Centered Multigrid for Interface Problems ». In : *Journal of Computational Physics* 98.1 (jan. 1992), p. 1-10. ISSN : 0021-9991. DOI : [10.1016/0021-9991\(92\)90168-X](https://doi.org/10.1016/0021-9991(92)90168-X). (Visité le 24/10/2022).
- [37] Kristin KIRCHNER et Christoph SCHWAB. *Monte Carlo Convergence Rates for Kth Moments in Banach Spaces*. Déc. 2022. arXiv : [2212.03797](https://arxiv.org/abs/2212.03797) [cs, math]. (Visité le 05/01/2023).
- [38] S. KRUMSCHEID et F. NOBILE. « Multilevel Monte Carlo Approximation of Functions ». In : *SIAM/ASA Journal on Uncertainty Quantification* 6.3 (jan. 2018), p. 1256-1293. DOI : [10.1137/17M1135566](https://doi.org/10.1137/17M1135566). (Visité le 19/02/2024).

- [39] Prashant KUMAR, Cornelis W. OOSTERLEE et Richard P. DWIGHT. « A Multigrid Multilevel Monte Carlo Method Using High-Order Finite-Volume Scheme for Lognormal Diffusion Problems ». In : *International Journal for Uncertainty Quantification* 7.1 (2017). ISSN : 2152-5080, 2152-5099. DOI : [10.1615/Int.J.UncertaintyQuantification.2016018677](https://doi.org/10.1615/Int.J.UncertaintyQuantification.2016018677). (Visité le 24/10/2022).
- [40] S. S LAVENBERG, T. L. MOELLER et P. D. WELCH. « The Application of Control Variables to the Simulation of Closed Queueing Networks ». In : *Proceedings of the 9th Conference on Winter Simulation - Volume 1. WSC '77*. Gaithersburg, Maryland, USA : Winter Simulation Conference, déc. 1977, p. 152-154. (Visité le 05/07/2024).
- [41] S. S. LAVENBERG et P. D. WELCH. « A Perspective on the Use of Control Variables to Increase the Efficiency of Monte Carlo Simulations ». In : *Management Science* 27.3 (mars 1981), p. 322-335. ISSN : 0025-1909. DOI : [10.1287/mnsc.27.3.322](https://doi.org/10.1287/mnsc.27.3.322). (Visité le 05/07/2024).
- [42] O. P. LE MAÎTRE et Omar M. KNIO. *Spectral Methods for Uncertainty Quantification : With Applications to Computational Fluid Dynamics*. Scientific Computation. Dordrecht : Springer Netherlands, 2010. ISBN : 978-90-481-3519-6 978-90-481-3520-2. DOI : [10.1007/978-90-481-3520-2](https://doi.org/10.1007/978-90-481-3520-2). (Visité le 23/03/2024).
- [43] Christiane LEMIEUX. « Control Variates ». In : *Wiley StatsRef : Statistics Reference Online*. John Wiley & Sons, Ltd, 2017, p. 1-8. ISBN : 978-1-118-44511-2. DOI : [10.1002/9781118445112.stat07947](https://doi.org/10.1002/9781118445112.stat07947). (Visité le 05/07/2024).
- [44] Christiane LEMIEUX. *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer Series in Statistics. New York, NY : Springer, 2009. ISBN : 978-0-387-78164-8 978-0-387-78165-5. DOI : [10.1007/978-0-387-78165-5](https://doi.org/10.1007/978-0-387-78165-5). (Visité le 28/06/2023).
- [45] Christopher LESTER et al. « Extending the Multi-level Method for the Simulation of Stochastic Biological Systems ». In : *Bulletin of Mathematical Biology* 78.8 (août 2016), p. 1640-1677. ISSN : 1522-9602. DOI : [10.1007/s11538-016-0178-9](https://doi.org/10.1007/s11538-016-0178-9). (Visité le 23/03/2024).
- [46] Finn LINDGREN, Håvard RUE et Johan LINDSTRÖM. « An Explicit Link between Gaussian Fields and Gaussian Markov Random Fields : The Stochastic Partial Differential Equation Approach ». In : *Journal of the Royal Statistical Society : Series B (Statistical Methodology)* 73.4 (2011), p. 423-498. ISSN : 1467-9868. DOI : [10.1111/j.1467-9868.2011.00777.x](https://doi.org/10.1111/j.1467-9868.2011.00777.x). (Visité le 26/02/2023).
- [47] Nora LÜTHEN, Stefano MARELLI et Bruno SUDRET. « Sparse Polynomial Chaos Expansions : Literature Survey and Benchmark ». In : *SIAM/ASA Journal on Uncertainty Quantification* 9.2 (jan. 2021), p. 593-649. DOI : [10.1137/20M1315774](https://doi.org/10.1137/20M1315774). (Visité le 23/03/2024).
- [48] Aimee MAURIS et al. « Multi-Fidelity Covariance Estimation in the Log-Euclidean Geometry ». In : *Proceedings of the 40th International Conference on Machine Learning*. PMLR, juill. 2023, p. 24214-24235. (Visité le 03/07/2024).
- [49] Aimee MAURIS et al. *Multifidelity Covariance Estimation via Regression on the Manifold of Symmetric Positive Definite Matrices*. Juill. 2023. DOI : [10.48550/arXiv.2307.12438](https://doi.org/10.48550/arXiv.2307.12438). arXiv : [2307.12438](https://arxiv.org/abs/2307.12438) [cs, math, stat]. (Visité le 03/07/2024).
- [50] Isabelle MIROUZE et Anthony T. WEAVER. « Representation of Correlation Functions in Variational Assimilation Using an Implicit Diffusion Operator ». In : *Quarterly Journal of the Royal Meteorological Society* 136.651 (2010), p. 1421-1443. ISSN : 1477-870X. DOI : [10.1002/qj.643](https://doi.org/10.1002/qj.643). (Visité le 04/01/2023).

- [51] Marcus MOHR et Roman WIENANDS. « Cell-Centred Multigrid Revisited ». In : *Computing and Visualization in Science* 7.3 (oct. 2004), p. 129-140. ISSN : 1433-0369. DOI : [10.1007/s00791-004-0137-0](https://doi.org/10.1007/s00791-004-0137-0). (Visité le 02/05/2024).
- [52] Paul MYCEK et Matthias DE LOZZO. « Multilevel Monte Carlo Covariance Estimation for the Computation of Sobol' Indices ». In : *SIAM/ASA Journal on Uncertainty Quantification* 7.4 (jan. 2019), p. 1323-1348. DOI : [10.1137/18M1216389](https://doi.org/10.1137/18M1216389). (Visité le 07/06/2022).
- [53] Habib N. NAJM. « Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics ». In : *Annual Review of Fluid Mechanics* 41. Volume 41, 2009 (jan. 2009), p. 35-52. ISSN : 0066-4189, 1545-4479. DOI : [10.1146/annurev.fluid.010908.165248](https://doi.org/10.1146/annurev.fluid.010908.165248). (Visité le 23/03/2024).
- [54] Barry L. NELSON. « On Control Variate Estimators ». In : *Computers & Operations Research* 14.3 (jan. 1987), p. 219-225. ISSN : 0305-0548. DOI : [10.1016/0305-0548\(87\)90024-4](https://doi.org/10.1016/0305-0548(87)90024-4). (Visité le 28/06/2023).
- [55] Leo W. T. NG et Karen E. WILLCOX. « Multifidelity Approaches for Optimization under Uncertainty ». In : *International Journal for Numerical Methods in Engineering* 100.10 (2014), p. 746-772. ISSN : 1097-0207. DOI : [10.1002/nme.4761](https://doi.org/10.1002/nme.4761). (Visité le 29/01/2024).
- [56] Benjamin PEHERSTORFER, Karen WILLCOX et Max GUNZBURGER. « Optimal Model Management for Multifidelity Monte Carlo Estimation ». In : *SIAM Journal on Scientific Computing* 38.5 (jan. 2016), A3163-A3194. ISSN : 1064-8275. DOI : [10.1137/15M1046472](https://doi.org/10.1137/15M1046472). (Visité le 25/06/2024).
- [57] Benjamin PEHERSTORFER, Karen WILLCOX et Max GUNZBURGER. « Survey of Multifidelity Methods in Uncertainty Propagation, Inference, and Optimization ». In : *SIAM Review* 60.3 (jan. 2018), p. 550-591. ISSN : 0036-1445. DOI : [10.1137/16M1082469](https://doi.org/10.1137/16M1082469). (Visité le 04/04/2023).
- [58] R. James PURSER et al. « Numerical Aspects of the Application of Recursive Filters to Variational Statistical Analysis. Part I : Spatially Homogeneous and Isotropic Gaussian Covariances ». In : *Monthly Weather Review* 131.8 (août 2003), p. 1524-1535. ISSN : 1520-0493, 0027-0644. DOI : [10.1175//1520-0493\(2003\)131<1524:NAOTA0>2.0.CO;2](https://doi.org/10.1175//1520-0493(2003)131<1524:NAOTA0>2.0.CO;2). (Visité le 23/03/2024).
- [59] Carl Edward RASMUSSEN. « Gaussian Processes in Machine Learning ». In : *Advanced Lectures on Machine Learning*. Sous la dir. d'Olivier BOUSQUET, Ulrike VON LUXBURG et Gunnar RÄTSCH. T. 3176. Berlin, Heidelberg : Springer Berlin Heidelberg, 2004, p. 63-71. ISBN : 978-3-540-23122-6 978-3-540-28650-9. DOI : [10.1007/978-3-540-28650-9\\_4](https://doi.org/10.1007/978-3-540-28650-9_4). (Visité le 23/03/2024).
- [60] Carl Edward RASMUSSEN et Christopher K. I. WILLIAMS. *Gaussian Processes for Machine Learning*. The MIT Press, nov. 2005. ISBN : 978-0-262-25683-4. DOI : [10.7551/mitpress/3206.001.0001](https://doi.org/10.7551/mitpress/3206.001.0001). (Visité le 23/03/2024).
- [61] Christian P. ROBERT et George CASELLA. *Monte Carlo Statistical Methods*. Springer Texts in Statistics. New York, NY : Springer, 2004. ISBN : 978-1-4419-1939-7 978-1-4757-4145-2. DOI : [10.1007/978-1-4757-4145-2](https://doi.org/10.1007/978-1-4757-4145-2). (Visité le 04/01/2023).
- [62] Daniel SCHADEN et Elisabeth ULLMANN. « Asymptotic Analysis of Multilevel Best Linear Unbiased Estimators ». In : *SIAM/ASA Journal on Uncertainty Quantification* 9.3 (jan. 2021), p. 953-978. DOI : [10.1137/20M1321607](https://doi.org/10.1137/20M1321607). (Visité le 04/01/2023).

- [63] Daniel SCHADEN et Elisabeth ULLMANN. « On Multilevel Best Linear Unbiased Estimators ». In : *SIAM/ASA Journal on Uncertainty Quantification* 8.2 (jan. 2020), p. 601-635. DOI : [10.1137/19M1263534](https://doi.org/10.1137/19M1263534). (Visité le 04/01/2023).
- [64] Christoph SCHWAB et Andreas STEIN. « Multilevel Monte Carlo FEM for Elliptic PDEs with Besov Random Tree Priors ». In : *Stochastics and Partial Differential Equations : Analysis and Computations* (sept. 2023). ISSN : 2194-041X. DOI : [10.1007/s40072-023-00313-w](https://doi.org/10.1007/s40072-023-00313-w). (Visité le 23/03/2024).
- [65] Ralph SHAPIRO. « Smoothing, Filtering, and Boundary Effects ». In : *Reviews of Geophysics* 8.2 (1970), p. 359-387. ISSN : 1944-9208. DOI : [10.1029/RG008i002p00359](https://doi.org/10.1029/RG008i002p00359). (Visité le 13/07/2022).
- [66] Jonas ŠUKYS et al. *Optimal Fidelity Multi-Level Monte Carlo for Quantification of Uncertainty in Simulations of Cloud Cavitation Collapse*. Mai 2017. DOI : [10.48550/arXiv.1705.04374](https://doi.org/10.48550/arXiv.1705.04374). arXiv : [1705.04374](https://arxiv.org/abs/1705.04374) [cs, math, stat]. (Visité le 25/06/2024).
- [67] Aretha L. TECKENTRUP et al. « Further Analysis of Multilevel Monte Carlo Methods for Elliptic PDEs with Random Coefficients ». In : *Numerische Mathematik* 125.3 (nov. 2013), p. 569-600. ISSN : 0945-3245. DOI : [10.1007/s00211-013-0546-4](https://doi.org/10.1007/s00211-013-0546-4). (Visité le 23/03/2024).
- [68] Rohit K. TRIPATHY et Ilias BILIONIS. « Deep UQ : Learning Deep Neural Network Surrogate Models for High Dimensional Uncertainty Quantification ». In : *Journal of Computational Physics* 375 (déc. 2018), p. 565-588. ISSN : 0021-9991. DOI : [10.1016/j.jcp.2018.08.036](https://doi.org/10.1016/j.jcp.2018.08.036). (Visité le 23/03/2024).
- [69] Ulrich TROTTEMBERG, Cornelius W. OOSTERLEE et Anton SCHULLER. *Multigrid*. Elsevier, nov. 2000. ISBN : 978-0-08-047956-9.
- [70] Anthony T. WEAVER et Philippe COURTIER. « Correlation modelling on the sphere using a generalized diffusion equation ». In : *Quarterly Journal of the Royal Meteorological Society* 127.575 (2001), p. 1815-1846. ISSN : 1477-870X. DOI : [10.1002/qj.49712757518](https://doi.org/10.1002/qj.49712757518). (Visité le 06/04/2023).
- [71] Anthony T. WEAVER et Isabelle MIROUZE. « On the Diffusion Equation and Its Application to Isotropic and Anisotropic Correlation Modelling in Variational Assimilation ». In : *Quarterly Journal of the Royal Meteorological Society* 139.670 (2013), p. 242-260. ISSN : 1477-870X. DOI : [10.1002/qj.1955](https://doi.org/10.1002/qj.1955). (Visité le 09/06/2023).
- [72] Anthony T. WEAVER, Jean TSHIMANGA et Andrea PIACENTINI. « Correlation Operators Based on an Implicitly Formulated Diffusion Equation Solved with the Chebyshev Iteration ». In : *Quarterly Journal of the Royal Meteorological Society* 142.694 (jan. 2016), p. 455-471. ISSN : 0035-9009, 1477-870X. DOI : [10.1002/qj.2664](https://doi.org/10.1002/qj.2664). (Visité le 08/04/2024).
- [73] Anthony T. WEAVER et al. « An Evaluation of Methods for Normalizing Diffusion-Based Covariance Operators in Variational Data Assimilation ». In : *Quarterly Journal of the Royal Meteorological Society* 147.734 (2021), p. 289-320. ISSN : 1477-870X. DOI : [10.1002/qj.3918](https://doi.org/10.1002/qj.3918). (Visité le 04/01/2023).
- [74] P WESSELING. « Cell-Centered Multigrid for Interface Problems ». In : *Journal of Computational Physics* 79.1 (nov. 1988), p. 85-91. ISSN : 0021-9991. DOI : [10.1016/0021-9991\(88\)90005-8](https://doi.org/10.1016/0021-9991(88)90005-8). (Visité le 31/07/2024).
- [75] Peter WHITTLE. « Stochastic-processes in several dimensions ». In : *Bulletin of the International Statistical Institute* 40.2 (1963), p. 974-994.

- 
- [76] Roman WIENANDS. *Practical Fourier Analysis for Multigrid Methods*. Chapman and Hall/CRC, oct. 2004. ISBN : 978-0-429-14328-1. DOI : [10 . 1201 / 9781420034998](https://doi.org/10.1201/9781420034998). (Visité le 04/01/2023).
- [77] Dongbin XIU. *Numerical Methods for Stochastic Computations : A Spectral Method Approach*. Princeton University Press, 2010. ISBN : 978-0-691-14212-8. DOI : [10 . 2307/j .ctv7h0skv](https://doi.org/10.2307/j.ctv7h0skv). JSTOR : [j .ctv7h0skv](https://www.jstor.org/stable/j.ctv7h0skv). (Visité le 23/03/2024).

**Titre :** Méthodes de Monte Carlo multi-niveaux pour l'estimation de paramètres statistiques de champs discrétisés en géosciences

**Mots clés :** Monte Carlo multi-niveaux, Multifidélité, Géosciences, Champ aléatoire, Analyse spectrale

**Résumé :** Les méthodes de réduction de variance multi-fidélités pour l'estimation statistique sont utilisées dans de plus en plus de domaines comme une amélioration par rapport à la méthode Monte Carlo. Dans leur forme originale, elles sont conçues pour estimer des paramètres statistiques scalaires. Cependant, dans certaines applications, par exemple en géosciences, les quantités d'intérêt dont on souhaite estimer les paramètres statistiques peuvent être des vecteurs ou des champs aléatoires. Dans le cas d'un vecteur aléatoire représentant un champ aléatoire discrétisé, l'utilisation d'un estimateur multi-fidélité n'est pas triviale, notamment si les entrées et sorties des différents modèles n'ont pas les mêmes dimensions sur les différents niveaux de fidélité. Des opérateurs de transfert entre les différents niveaux doivent alors être introduits. Cette thèse porte sur l'adaptation de l'estimateur Monte Carlo multi-niveaux (MLMC) à l'estimation de champs discrétisés en géosciences et sur l'analyse des opérateurs de transfert, permettant le choix d'opérateurs qui réduisent la variance du MLMC.

Afin de mieux comprendre l'effet des opérateurs de transfert, une analyse spectrale de l'estimateur MLMC est proposée. Des expériences numériques sur des problèmes simplifiés ont mis en évidence que le MLMC détériore l'estimation des composantes hautes fréquences d'un champ discrétisé comparé à un simple estimateur Monte Carlo. Une analyse théorique portant sur une classe de problèmes spécifique, similaire aux analyses développées pour les méthodes multigrilles, a permis une meilleure compréhension des disparités dans l'estimation, par MLMC, des différentes composantes d'un champ discrétisé. Ces différents résultats nous ont conduit à modifier les opérateurs de transfert utilisés afin d'y ajouter une étape de filtrage inspirée des méthodes multigrilles. L'ajout de filtres permet d'améliorer l'estimation des composantes hautes et basses fréquences, réduisant ainsi la variance totale de l'estimateur.

L'utilisation de méthodes de type multilevel best linear unbiased estimator (MLBLUE) a permis de mieux choisir les opérateurs de transfert et les filtres en vue de réduire la variance de l'estimateur. Ces différentes améliorations ont été appliquées à l'estimation de la variance d'un champ discrétisé obtenu par application d'un opérateur de covariance basé sur la diffusion. Les expériences numériques ont pu confirmer les résultats théoriques sur un problème plus complexe et mettre en évidence les améliorations apportées par rapport à un simple MLMC.

Cette recherche contribue à élargir les applications du MLMC à d'autres domaines, notamment les géosciences, et offre une meilleure compréhension de la méthode grâce aux liens tissés avec les méthodes multigrilles. Elle permet également d'améliorer l'estimateur MLMC en soulignant l'importance du choix des opérateurs de transfert.

**Title:** Multilevel Monte Carlo methods for the estimation of statistics of discretized fields in geosciences

**Key words:** Multilevel Monte Carlo, Multifidelity, Geosciences, Random field, Spectral analysis

**Abstract:** Variance reduction multifidelity methods for the estimation of statistics are increasingly used in more domains as an improvement over the classical Monte Carlo method. These methods were originally designed to estimate scalar statistics. However, in some applications, in geosciences for example, the quantities of interest which statistics we want to estimate can be random vectors or random fields. In the case of a random vector which represent a discretized field, using a multifidelity estimator is not straightforward, especially if the inputs and outputs of the different models do not have the same dimension among the different fidelity levels. Transfer operators between the levels need to be introduced.

This thesis is about the adaptation of the multilevel Monte Carlo (MLMC) estimator for the estimation of discretized fields in geosciences and the analysis of the transfer operators, allowing to choose operators that reduce the variance of the MLMC.

A spectral analysis of the MLMC estimator is performed to better understand the effects of the transfer operators. Numerical experiments on simplified problems show that the MLMC deteriorates the estimation of high frequency components of a discretized field compared to a Monte Carlo estimator. A theoretical analysis on a specific class of problems, similar to analysis developed for multigrid methods, allows for a better understanding of the disparities in the MLMC estimation of the different components of a discretized field. Following those results, the transfer operators are modified to include a filtering step inspired from multigrid methods. Adding filters leads to a better estimation of low and high frequency components, and to a decrease in the total variance of the estimator.

The use of methods similar to the multilevel best linear unbiased estimator (MLBLUE) can help in choosing the transfer operators that reduce the variance of the estimator. Those improvements were applied to the estimation of the discretized variance field of a diffusion-based covariance operator. Numerical experiments support the conclusions of the theoretical analysis on a more complex problem and demonstrates the improvements brought by the proposed estimator compared to a crude MLMC.

This research work contributes to expand the range of MLMC applications to other domains, especially geosciences, and offer a better understanding of the method by establishing links with multigrid methods. It also allows for improvements of the MLMC estimator by highlighting the importance of the transfer operator choice.