



HAL
open science

Sur l'estimation de paramètres dans les modèles différentiels stochastiques multidimensionnels

Alain Le Breton

► **To cite this version:**

Alain Le Breton. Sur l'estimation de paramètres dans les modèles différentiels stochastiques multidimensionnels. Modélisation et simulation. Institut National Polytechnique de Grenoble - INPG; Université Joseph-Fourier - Grenoble I, 1976. tel-00287156

HAL Id: tel-00287156

<https://theses.hal.science/tel-00287156>

Submitted on 11 Jun 2008

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.

Monsieur Gabriel CAU : Président
Monsieur Pierre JULLIEN : Vice-Président

MEMBRES DU CORPS ENSEIGNANT DE L'U.S.M.G.

PROFESSEURS TITULAIRES

MM. ARNAUD Paul	Chimie
AUBERT Guy	Physique
AYANT Yves	Physique approfondie
Mme BARBIER Marie-Jeanne	Electrochimie
MM. BARBIER Jean-Claude	Physique Expérimentale
BARBIER Reynold	Géologie appliquée
BARJON Robert	Physique nucléaire
BARNOUD Fernand	Biosynthèse de la cellulose
BARRA Jean-René	Statistiques
BARRIE Joseph	Clinique chirurgicale
BEAUDOING André	Clinique de Pédiatrie et Puériculture
BERNARD Alain	Mathématiques Pures
Mme BERTRANDIAS Françoise	Mathématiques Pures
MM. BERTRANDIAS Jean-Paul	Mathématiques Pures
BEZES Henri	Pathologie chirurgicale
BLAMBERT Maurice	Mathématiques Pures
BOLLIET Louis	Informatique (IUT B)
BONNET Georges	Electrotechnique
BONNET Jean-Louis	Clinique ophtalmologique
BONNET-EYMARD Joseph	Clinique gastro-entérologique
Mme BONNIER Marie-Jeanne	Chimie générale
MM. BOUCHERLE André	Chimie et toxicologie
BOUCHEZ Robert	Physique nucléaire
BOUSSARD Jean-Claude	Mathématiques Appliquées
BOUTET DE MONTVEL Louis	Mathématiques Pures
BRAVARD Yves	Géographie
CABANEL Guy	Clinique rhumatologique et hydrologique
CALAS François	Anatomie
CARLIER Georges	Biologie végétale
CARRAZ Gilbert	Biologie animale et pharmacodynamie
CAU Gabriel	Médecine légale et toxicologie
CAUQUIS Georges	Chimie organique
CHABAUTY Claude	Mathématiques Pures
CHARACHON Robert	Clinique Oto-rhino-laryngologique
CHATEAU Robert	Clinique de neurologie
CHIBON Pierre	Biologie animale
COEUR André	Pharmacie chimique et chimie analytique
CONTAMIN Robert	Clinique gynécologique
COUDERC Pierre	Anatomie pathologique
Mme DEBELMAS Anne-Marie	Matière médicale
MM. DEBELMAS Jacques	Géologie générale
DEGRANGE Charles	Zoologie
DELORMAS Pierre	Pneumophtisiologie

MM. DEPORTES Charles	Chimie minérale
DESRE Pierre	Métallurgie
DESSAUX Georges	Physiologie animale
DODU Jacques	Mécanique appliquée (IUT A)
DOLIQUE Jean-Michel	Physique des plasmas
DREYFUS Bernard	Thermodynamique
DUCROS Pierre	Cristallographie
DUGOIS Pierre	Clinique de dermatologie et syphiligraphie
GAGNAIRE Didier	Chimie physique
GALLISSOT François	Mathématiques Pures
GALVANI Octave	Mathématiques Pures
GASTINEL Noël	Analyse numérique
GAVEND Michel	Pharmacologie
GEINDRE Michel	Electroradiologie
GERBER Robert	Mathématiques Pures
GERMAIN Jean-Pierre	Mécanique
GIRAUD Pierre	Géologie
JANIN Bernard	Géographie
KAHANE André	Physique générale
KLEIN Joseph	Mathématiques Pures
KOSZUL Jean-Louis	Mathématiques Pures
KRAVTCHENKO Julien	Mécanique
KUNTZMANN Jean	Mathématiques Appliquées
LACAZE Albert	Thermodynamique
LACHARME Jean	Biologie végétale
Mme LAJZEROWICZ Janine	Physique
MM. LAJZEROWICZ Joseph	Physique
LATREILLE René	Chirurgie générale
LATURAZE Jean	Biochimie pharmaceutique
LAURENT Pierre-Jean	Mathématiques Appliquées
LEDRU Jean	Clinique médicale B
LLIBOUTRY Louis	Géophysique
LOISEAUX Pierre	Sciences nucléaires
LONGEQUEUE Jean-Pierre	Physique nucléaire
LOUP Jean	Géographie
Mlle LUTZ Elisabeth	Mathématiques Pures
MM. MALGRANGE Bernard	Mathématiques Pures
MALINAS Yves	Clinique obstétricale
MARTIN-NOEL Pierre	Clinique cardiologique
MAZARE Yves	Clinique médicale A
MICHEL Robert	Minéralogie et Pétrographie
MICLOUD Max	Clinique maladies infectieuses
MOURIQUAND Claude	Histologie
MOUSSA André	Chimie nucléaire
MULLER Jean-Michel	Thérapeutique (Néphrologie)
NEEL Louis	Physique du Solide
OZENDA Paul	Botanique
PAYAN Jean-Jacques	Mathématiques Pures
PEBAY-PEYROULA Jean-Claude	Physique
RASSAT André	Chimie systématique
RENARD Michel	Thermodynamique
REVOL Michel	Urologie
RINALDI Renaud	Physique
DE ROUGEMONT Jacques	Neuro-chirurgie
SEIGNEURIN Raymond	Microbiologie et Hygiène
SENGEL Philippe	Zoologie

MM. SIBILLE Robert	Construction mécanique (IUT A)
SOUTIF Michel	Physique générale
TANCHE Maurice	Physiologie
TRAYNARD Philippe	Chimie générale
VAILLANT François	Zoologie
VALENTIN Jacques	Physique nucléaire
VAUQUOIS Bernard	Calcul électronique
Mme VERAÏN Alice	Pharmacie galénique
MM. VERAÏN André	Physique
VEYRET Paul	Géographie
VIGNAIS Pierre	Biochimie médicale
YOCCOZ Jean	Physique nucléaire théorique

PROFESSEURS ASSOCIES

MM. CLARK Gilbert	Spectrométrie physique
CRABBE Pierre	CERMO
ENGLMAN Robert	Spectrométrie physique
HOLTZBERG Frédéric	Basses températures
DEMBICKI Eugéniuz	Mécanique
MATSUSHIMA Yozo	Mathématiques Pures

PROFESSEURS SANS CHAIRE

Mlle AGNIUS-DELORD Claudine	Physique pharmaceutique
ALARY Josette	Chimie analytique
MM. AMBROISE-THOMAS Pierre	Parasitologie
BELORIZKY Elie	Physique
BENZAKEN Claude	Mathématiques Appliquées
BIAREZ Jean-Pierre	Mécanique
BILLET Jean	Géographie
BOUCHET Yves	Anatomie
BRUGEL Lucien	Energétique (IUT A)
BUISSON René	Physique (IUT A)
BUTEL Jean	Orthopédie
COHEN ADDAD Pierre	Spectrométrie physique
COLOMB Maurice	Biochimie
CONTE René	Physique (IUT A)
DEPASSEL Roger	Mécanique des fluides
FONTAINE Jean-Marc	Mathématiques Pures
GAUTHIER Yves	Sciences Biologiques
GAUTRON René	Chimie
GIDON Paul	Géologie et Minéralogie
GLENAT René	Chimie organique
GROULADE Joseph	Biochimie médicale
HACQUES Gérard	Calcul numérique
HOLLARD Daniel	Hématologie
HUGONOT Robert	Hygiène et Médecine préventive
IDELMAN Simon	Physiologie animale
JOLY Jean-René	Mathématiques Pures
JULLIEN Pierre	Mathématiques Appliquées
Mme KAHANE Josette	Physique
MM. KRAKOWIAK Sacha	Mathématiques Appliquées
KUHN Gérard	Physique (IUT A)
LE ROY Philippe	Mécanique (IUT A)
LJU DUC Cuong	Chimie organique

MM. MAYNARD Roger	Physique du solide
Mme MINIER Colette	Physique (IUT A)
MM. PELMONT Jean	Biochimie
PERRIAUX Jean-Jacques	Géologie et Minéralogie
PFISTER Jean-Claude	Physique du solide
Mlle PIERY Yvette	Physiologie animale
MM. RAYNAUD Hervé	M.I.A.G.
REBECQ Jacques	Biologie (CUS)
REYMOND Jean-Charles	Chirurgie générale
RICHARD Lucien	Biologie végétale
Mme RINAUDO Marguerite	Chimie macromoléculaire
MM. ROBERT André	Chimie papetière
SARRAZIN Roger	Anatomie et chirurgie
SARROT-REYNAULD Jean	Géologie
SIROT Louis	Chirurgie générale
Mme SOUTIF Jeanne	Physique générale
MM. STREGLITZ Paul	Anesthésiologie
VIALON Pierre	Géologie
VAN CUTSEM Bernard	Mathématiques Appliquées

MATTRES DE CONFERENCES ET MATTRES DE CONFERENCES AGREGES

MM. AMBLARD Pierre	Dermatologie
ARMAND Gilbert	Géographie
ARMAND Yves	Chimie (IUT A)
BACHELOT Yvan	Endocrinologie
BARGE Michel	Neuro-chirurgie
BARJOLLE Michel	M.I.A.G.
BEGUIN Claude	Chimie organique
Mme BERIEL Hélène	Pharmacodynamie
MM. BOST Michel	Pédiatrie
BOUCHARLAT Jacques	Psychiatrie adultes
Mme BOUCHE Liane	Mathématiques (CUS)
MM. BRODEAU François	Mathématiques (IUT B)
CHAMBAZ Edmond	Biochimie médicale
CHAMPETIER Jean	Anatomie et organogénèse
CHARDON Michel	Géographie
CHERADAME Hervé	Chimie papetière
CHIAVERINA Jean	Biologie appliquée (EFP)
CONTAMIN Charles	Chirurgie thoracique et cardio-vasculaire
CORDONNIER Daniel	Néphrologie
COULOMB Max	Radiologie
CROUZET Guy	Radiologie
CYROT Michel	Physique du solide
DELOBEL Claude	M.I.A.G.
DENIS Bernard	Cardiologie
DOUCE Roland	Physiologie végétale
DUSSAUD René	Mathématiques (CUS)
Mme ETERRADOSSI Jacqueline	Physiologie
MM. FAURE Jacques	Médecine légale
FAURE Gilbert	Urologie
GAUTIER Robert	Chirurgie générale
GENSAC Pierre	Botanique
GIDON Maurice	Géologie
GROS Yves	Physique (IUT A)

MM. GUITTON Jacques	Chimie
HICTER Pierre	Chimie
IVANES Marcel	Electricité
JALBERT Pierre	Histologie
JUNIEN-LAVILLAVROY Claude	O.R.L.
KOLODIE Lucien	Hématologie
LE NOC Pierre	Bactériologie-virologie
LEROY Philippe	IUT A
MACHE Régis	Physiologie végétale
MAGNIN Robert	Hygiène et médecine préventive
MALLION Jean-Michel	Médecine du travail
MARECHAL Jean	Mécanique (IUT A)
MARTIN-BOUYER Michel	Chimie (CUS)
MICHOULIER Jean	Physique (IUT A)
NEGRE Robert	Mécanique (IUT A)
NEMOZ Alain	Thermodynamique
NOUGARET Marcel	Automatique (IUT A)
PARAMELLE Bernard	Pneumologie
PECCOUD François	Analyse (IUT B)
PEFFEN René	Métallurgie (IUT A)
PERRET Jean	Neurologie
PERRIER Guy	Géophysique - Glaciologie
PHELIP Xavier	Rhumatologie
RACHAIL Michel	Médecine interne
RACINET Claude	Gynécologie et obstétrique
RAMBAUD André	Hygiène et hydrologie
RAMBAUD Pierre	Pédiatrie
Mme RENAUDET Jacqueline	Bactériologie
MM. ROBERT Jean-Bernard	Chimie Physique
ROMIER Guy	Mathématiques (IUT B)
SHOM Jean-Claude	Chimie générale
STOEBNER Pierre	Anatomie pathologique
VROUSOS Constantin	Radiologie

MAITRE DE CONFERENCES ASSOCIES

M. COLE Antony

Sciences nucléaires

Fait à SAINT MARTIN D'HERES, AVRIL 1976.

INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE

Président : M. Philippe TRAYNARD
 Vice-Président : M. Pierre-Jean LAURENT

PROFESSEURS TITULAIRES

MM. BENOIT Jean	Radioélectricité
BESSON Jean	Electrochimie
BLOCH Daniel	Physique du solide
BONNETAIN Lucien	Chimie Minérale
BONNIER Etienne	Electrochimie et Electrometallurgie
BOUDOURIS Georges	Radioélectricité
BRISSONNEAU Pierre	Physique du solide
BUYLE-BODIN Maurice	Electronique
COUMES André	Radioélectricité
DURAND Francis	Métallurgie
FELICI Noël	Electrostatique
FOULARD Claude	Automatique
LESPINARD Georges	Mécanique
MOREAU René	Mécanique
PARIAUD Jean-Charles	Chimie-Physique
PAUTHENET René	Physique du solide
PERRET René	Servomécanismes
POLOJADOFF Michel	Electrotechnique
SILBER Robert	Mécanique des Fluides

PROFESSEUR ASSOCIE

M. ROUXEL Roland Automatique

PROFESSEURS SANS CHAIRE

MM. BLIMAN Samuel	Electronique
BOUVARD Maurice	Génie Mécanique
COHEN Joseph	Electrotechnique
LACOUME Jean-Louis	Géophysique
LANCIA Roland	Electronique
ROBERT François	Analyse numérique
VEILLON Gérard	Informatique Fondamentale et Appliquée
ZADWORNY François	Electronique

MAITRES DE CONFERENCES

MM. ANCEAU François	Mathématiques Appliquées
CHARTIER Germain	Electronique
GUYOT Pierre	Chimie Minérale
IVANES Marcel	Electrotechnique
JOUBERT Jean-Claude	Physique du solide
MORET Roger	Electrotechnique Nucléaire
PIERRARD Jean-Marie	Mécanique
SABONNADIÈRE Jean-Claude	Informatique Fondamentale et Appliquée
Mme SAUCIER Gabrièle	Informatique Fondamentale et Appliquée

MAITRE DE CONFERENCES ASSOCIE

M. LANDAU Ioan

Automatique

CHERCHEURS DU C.N.R.S. (Directeur et Maître de Recherche)

MM. FRUCHART Robert

Directeur de Recherche

ANSARA Ibrahim

Maître de Recherche

CARRE René

Maître de Recherche

DRIOLE Jean

Maître de Recherche

MATHIEU Jean-Claude

Maître de Recherche

MUNIER Jacques

Maître de Recherche

Je tiens à remercier ici tous ceux qui, tant par leur enseignement, leurs conseils et critiques, que par leurs témoignages d'intérêt et d'amitié, ont contribué à la réalisation de cette thèse.

Monsieur J.R. BARRA, Professeur à l'Université Scientifique et Médicale de Grenoble, m'a accueilli au sein de l'Equipe de Statistique où cette thèse a été élaborée. Je le remercie profondément de l'honneur qu'il me fait en acceptant de présider le Jury.

Monsieur F. BRODEAU, Professeur à l'Université des Sciences Sociales de Grenoble, a dirigé ce travail. Je le remercie très sincèrement de m'avoir toujours accordé sa confiance et d'avoir, par ses conseils et encouragements constants, su favoriser l'élaboration de cette thèse.

A plusieurs occasions j'ai profité des enseignements de Monsieur M. METIVIER, Professeur à l'Université de Rennes, qui ont suscité et entretenu mon goût pour la théorie des probabilités. Je suis particulièrement sensible à l'attention qu'il a portée à ce travail et à l'honneur qu'il me fait en acceptant de participer au Jury.

Messieurs U.G. HAUSSMANN, Professeur à l'Université de Colombie Britannique, et B. VAN CUTSEM, Professeur à l'Université Scientifique et Médicale de Grenoble, ont accepté de prendre part au Jury. Je les en remercie très vivement.

Mademoiselle MARCHAND a apporté beaucoup de soin à la frappe du manuscrit et les personnels des services de reprographie de l'I.R.M.A. et de l'I.M.S.S. ont réalisé un tirage de qualité. Je les remercie sincèrement et je tiens à saluer, ici, leur goût pour le travail bien fait.

TABLE DES MATIERES

	pages
INTRODUCTION	0 - 1
CHAPITRE I - LE MODELE GENERAL - LE PROBLEME - LES OUTILS.	I - 1
§ 1. Solutions faible et forte d'une équation différentielle stochastique. Position du problème d'estimation.	I - 2
1.1. Notations et définitions.	I - 2
1.2. Théorèmes d'équivalence et d'absolue continuité de lois sur (C, S_t)	I - 5
1.3. Existences et unicités forte et faible de solutions.	I - 9
1.4. Position du problème d'estimation. Notions d'identi- fiabilité.	I - 12
§ 2. Estimation par la méthode du maximum de vraisem- blance lorsque l'observation porte sur une fonction aléatoire en temps continu.	I - 15
2.1. Quelques lemmes.	I - 15
2.2. Cas de l'observation en temps continu.	I - 18
2.3. Cas de l'observation en temps discret.	I - 27
§ 3. Quelques propriétés des intégrales stochastiques. Dérivabilité par rapport à un paramètre. Normalité asymptotique.	I - 37
3.1. Dérivabilité par rapport à un paramètre.	I - 37
3.2. Normalité asymptotique.	I - 42
CHAPITRE II - ESTIMATION DANS LES MODELES DIFFERENTIELS STOCHASTIQUES A COEFFICIENTS DEPENDANT LINEAIREMENT DES PARAMETRES INCONNUS.	II - 1
§ 1. Estimation dans le modèle linéaire. Cas de l'obser- vation en temps continu.	II - 2
1.1. Formulation directe du problème. Structure statistique.	II - 3
1.2. Estimation de la matrice $G \cdot tG$.	II - 6
1.3. Estimation de la matrice θ .	II - 8
§ 2. Estimation dans le modèle linéaire. Cas de l'observa- tion en temps discret.	II - 25

2.1.	La fonction de Log-vraisemblance pour l'estimation basée sur un échantillonnage discret.	II - 26
2.2.	La fonction de Log-vraisemblance approchée pour l'estimation basée sur un échantillonnage discret.	II - 30
2.3.	La fonction de Log-vraisemblance approchée pour l'estimation basée sur un échantillonnage continu.	II - 38
2.4.	Estimation du paramètre de dérive.	II - 42
§ 3.	Estimation dans les modèles non linéaires à paramètres multiplicatifs dans les coefficients de diffusion et de dérive.	II - 44
3.1.	Estimation des paramètres du coefficient de diffusion.	II - 45
3.2.	Estimation des paramètres de dérive par maximum de vraisemblance.	II - 48
3.3.	Estimation "séquentielle" de maximum de vraisemblance des paramètres de dérive.	II - 54
3.4.	Estimation bayésienne des paramètres de dérive.	II - 58
3.5.	Quelques résultats numériques.	II - 60

CHAPITRE III - ESTIMATION DANS LES MODELES DIFFERENTIELS STOCHASTIQUES A COEFFICIENTS DEPENDANT NON LINEAIREMENT DES PARAMETRES INCONNUS.

§ 1.	Estimation dans un modèle différentiel stochastique à coefficients dépendant différentiablement des paramètres inconnus.	III - 1
1.1.	Estimation des paramètres du coefficient de diffusion.	III - 2
1.2.	Estimation des paramètres du coefficient de dérive.	III - 3
1.3.	Sur la construction des estimateurs.	III - 7
§ 2.	Application. Estimation de paramètres dans le modèle de filtrage linéaire à coefficients constants.	III - 15
2.1.	La structure statistique du problème d'estimation.	III - 17
2.2.	Estimation de la covariance du bruit d'observation.	III - 19
2.3.	Estimation des paramètres ne figurant pas dans la covariance du bruit d'observation.	III - 26
2.4.	Sur la condition d'identifiabilité et le calcul de l'estimateur de ϕ .	III - 27
§ 3.	Sur la modélisation et l'identification des systèmes stochastiques à temps continu. Quelques remarques pour conclure.	III - 48
3.1.	Sur l'identification des dimensions.	III - 50
3.2.	Sur le choix de l'intégrale stochastique et la nature du bruit.	III - 51
		III - 52

3.3. Sur l'identification : estimations, domaines de confiance,
tests.

III - 53

BIBLIOGRAPHIE

B - 1

-:-:-

N.B. Lorsque, dans un chapitre, il est fait référence à un énoncé d'un autre chapitre, le numéro de cet énoncé est précédé du chiffre romain correspondant à ce chapitre. Il en est de même pour les références à certains paragraphes.

INTRODUCTION

Le travail que nous présentons ici concerne une classe particulière de modèles de systèmes dynamiques stochastiques évoluant en temps continu : celle des modèles décrits par des équations différentielles stochastiques vectorielles de dimension finie. Il traite du problème statistique d'estimation de paramètres dans de tels modèles et, partant, s'inscrit dans le cadre de la statistique mathématique.

Exposons d'abord brièvement quelle était notre motivation et quels étaient nos buts en commençant ce travail. De nombreux phénomènes évolutifs peuvent être représentés de façon satisfaisante par des processus stochastiques de type diffusion dont la modélisation par une équation d'évolution fait intervenir des équations différentielles stochastiques. On en trouve des exemples concrets dans des domaines très divers tels que ceux de la météorologie, de l'aéronautique, de la physique ... De même, dans de nombreux problèmes de contrôle stochastique, l'évolution de l'état du système à contrôler ainsi que son observation peuvent raisonnablement être modélisées par des équations différentielles stochastiques. L'étude qualitative des processus mis en jeu et leur éventuel contrôle dépendent de façon cruciale de la plus ou moins bonne connaissance que l'on a des fonctions ou paramètres inconnus qui interviennent dans les équations de la modélisation. En pratique, l'estimation des paramètres inconnus doit généralement pouvoir être effectuée en temps réel, ce qui exclut la possibilité de répéter l'expérience fournissant l'observation sur laquelle est basée l'estimation. Nous nous sommes alors proposés d'étudier le problème d'estimation de paramètres inconnus de dimension finie dans un modèle différentiel stochastique, au vu de l'observation, en temps continu ou en temps discret, d'une seule trajectoire du processus qu'il engendre. Notre objectif était de

dégager des procédures d'estimation, d'en décrire de façon précise les qualités et d'en éprouver l'efficacité par des expériences de simulation. Décrivons maintenant comment nous avons organisé la présentation de notre travail.

Le premier chapitre est consacré à la formulation du problème général et à la présentation des outils que nous nous proposons d'utiliser pour en aborder l'étude.

Dans une première partie, nous décrivons la classe des modèles objet de notre travail en précisant le concept que nous utilisons de solution d'une équation différentielle stochastique. Il importe au statisticien de définir en premier lieu la structure statistique associée aux observations dont il dispose, c'est-à-dire essentiellement de décrire la famille des lois de probabilité possibles sur l'espace où ces observations prennent leurs valeurs. Dans cet esprit il est naturel de décrire les processus stochastiques mis en jeu par leurs lois de probabilité et par suite de choisir de définir la solution d'une équation différentielle stochastique comme étant une loi de probabilité sur un espace de trajectoires. Nous adoptons alors le formalisme de la statistique mathématique, tel qu'il a été introduit par J.R. BARRA dans son livre [9], pour formuler la problème d'estimation des paramètres inconnus dont dépendent les coefficients d'une équation différentielle stochastique. Les théorèmes d'existence d'une solution faible d'une équation différentielle stochastique reposent essentiellement sur un théorème de I.V. GIRSANOV [24] qui permet de définir une telle solution par sa dérivée de Radon-Nikodym par rapport à une mesure de référence. Ainsi, lorsque cette mesure de référence ne dépend pas de paramètres inconnus (il s'agira en fait du cas où le coefficient de diffusion du modèle est complètement connu), la structure statistique concernée est dominée, ce qui suggère l'approche du problème d'estimation par la méthode du maximum de vraisemblance.

Aussi, dans une deuxième partie, nous proposons des énoncés généraux sur l'estimation par maximum de vraisemblance dans une structure

statistique associée à l'observation d'une fonction aléatoire en temps continu que cette observation soit elle-même de dimension infinie (observation en temps continu) ou de dimension finie (observation en temps discret). Pour le cas de l'observation en temps continu nous pensons avoir dégagé un cadre d'hypothèses minimal permettant d'assurer l'existence d'une famille convergente et asymptotiquement normale d'estimateurs. Pour le cas de l'observation en temps discret, nous étudions la possibilité d'approcher l'estimateur basé sur un échantillonnage continu par l'estimateur construit sur un échantillonnage discret.

En vue de l'application de ces énoncés dans le contexte qui nous occupe, nous consacrons une troisième partie à l'exposé de quelques propriétés utiles des intégrales stochastiques. En premier lieu, afin d'assurer l'existence de versions régulières de fonctions de vraisemblance, nous étudions, pour une intégrale stochastique dont l'intégrande dépend d'un paramètre vectoriel, le problème de l'existence d'une version presque sûrement à trajectoires dérivables par rapport au paramètre du processus indexé par ce paramètre. Dans [23] I.I. GIHMAN et A.V. SKOROHOD donnent seulement des propriétés de dérivabilité en moyenne quadratique par rapport à un paramètre scalaire. Enfin, dans le but d'assurer la normalité asymptotique d'estimateurs, nous donnons un résultat sur la normalité asymptotique d'une intégrale stochastique que nous rapprochons d'énoncés de A.F. TARASKIN [49], [50].

Le second chapitre traite le problème d'estimation dans le cas des modèles différentiels dont les coefficients dépendent linéairement des paramètres inconnus. Dans les deux premières parties de ce chapitre, nous étudions le cas plus particulier où l'équation régissant l'observation est aussi linéaire en l'état du processus observé.

Dans la première partie, nous considérons la situation où l'observation a lieu en temps continu. A notre connaissance, tous les auteurs ayant abordé avant nous le problème d'estimation dans ce type de modèles

soit ont, en vue d'appliquer la méthode de maximum de vraisemblance, supposé connue la covariance du bruit d'observation sans donner de méthode d'estimation préalable de cette covariance (A.A. NOVIKOV [41], A.N. SHIRYAYEV [47] pour des modèles scalaires et bidimensionnels particuliers, ...), soit on proposé des techniques d'identification simultanée de tous les coefficients, sous l'hypothèse de stationnarité, utilisant essentiellement la fonction d'autocorrélation de l'observation (R.S. BUCY et P.D. JOSEPH [14]...), les propriétés des estimateurs restant mal connues. Nous montrons ici qu'il est possible de découpler l'estimation du coefficient de diffusion et l'estimation du coefficient de dérive. De façon précise, nous avons remarqué que la covariance du bruit d'observation peut être évaluée, avec probabilité un, sur tout intervalle de temps fini et que donc il est légitime de la supposer connue pour l'estimation du coefficient de dérive. Nous précisons aussi les qualités de la suite d'estimateurs envisagée. Alors la méthode du maximum de vraisemblance s'applique pour le problème d'estimation de la matrice de dérive et fournit un estimateur dont nous montrons que, sous certaines hypothèses, il est convergent en probabilité et asymptotiquement normalement distribué. Les résultats de cette partie ont été annoncés dans [30] A. LE BRETON, et développés dans [31] A. LE BRETON. L'estimateur de maximum de vraisemblance de la matrice de dérive basé sur l'observation en temps continu est défini en termes d'intégrales stochastique et ordinaire. Aussi, en pratique, pour le calculer, on peut penser à approximer ces intégrales par des sommes finies appropriées qui dépendent seulement d'un échantillonnage effectué en temps discret. Une question naturelle se pose alors : l'estimateur obtenu par cette discrétisation est-il "le meilleur" que l'on aurait pu obtenir au vu de l'échantillonnage discret qu'il met en jeu ? Nous sommes donc conduits à traiter le problème d'estimation lorsque l'observation est faite en temps discret.

C'est l'objet de la deuxième partie de ce chapitre où nous répondons en particulier à la question précédente en montrant, grâce à un énoncé du premier chapitre, qu'une "bonne" approximation de l'estimateur

de maximum de vraisemblance basée sur un échantillonnage discret est précisément l'estimateur résultant de la discrétisation de l'estimateur obtenu en temps continu qui fait apparaître cet échantillonnage discret. Les résultats de cette partie ont été présentés dans [33] A. LE BRETON.

Dans la troisième partie de ce chapitre, nous présentons d'abord une généralisation des développements de la première partie dans le cas des modèles vectoriels non linéaires en l'état, mais où toutefois les paramètres inconnus apparaissent de manière multiplicative (donc linéaire) dans les coefficients de dérive et de diffusion de l'équation différentielle stochastique concernée, modèles du type de celui considéré dans [49] par A.F. TARASKIN qui lui aussi suppose complètement connu le coefficient de diffusion. Puis nous envisageons la construction d'autres estimations pour de tels modèles : estimation de maximum de vraisemblance séquentielle selon une technique proposée pour le cas scalaire par P.S. LIPCER et A.N. SHIRYAYEV dans [36], et estimation bayésienne. Ce chapitre se termine sur la présentation de quelques résultats numériques d'expériences de simulation qui nous semblent encourager à l'utilisation des méthodes que nous avons étudiées pour la modélisation de diffusions stochastiques.

Le troisième chapitre traite le problème d'estimation dans le cas des modèles différentiels stochastiques où les coefficients de l'équation régissant l'observation dépendent de façon non linéaire des paramètres inconnus.

Dans une première partie, nous envisageons le cas d'un modèle général où les coefficients dépendent différenciablement des paramètres. Nous montrons que, sous certaines hypothèses, il est encore possible de découpler l'estimation des paramètres apparaissant dans le coefficient de diffusion et celle des autres paramètres. Le terme de diffusion étant alors supposé complètement connu, nous assurons l'existence d'une famille convergente et asymptotiquement normale d'estimateurs des paramètres restant à estimer. Les résultats, qui ont été présentés dans [34] A. LE BRETON,

sont de nature plus qualitative que ceux du deuxième chapitre, car nous n'avons pas ici d'expressions explicites des estimateurs. Toutefois l'utilisation pratique des résultats est suggérée par les énoncés eux-mêmes qui montrent, aussi bien dans le cas des paramètres de diffusion que dans celui des paramètres de dérive, que la construction des estimateurs repose sur la mise en oeuvre de méthodes numériques de résolution d'une équation vectorielle à coefficients dépendant de la trajectoire observée du processus ; nous indiquons une approche algorithmique de ce problème.

Dans une deuxième partie, nous étudions le cas particulier du modèle classique de la théorie du filtrage linéaire en temps continu. Pour un tel modèle, on dispose, sans supposer connue la dimension de l'état non observable, mais sous l'hypothèse de stationnarité, de techniques d'identification telles que celle proposée par B.L. HO et R.E. KALMAN [26]. Supposant connue la dimension de l'état non observable, éventuellement parce qu'une méthode du type précédent aurait permis de l'identifier, en utilisant la représentation du modèle en termes de l'innovation, nous montrons que, bien que les paramètres inconnus apparaissent linéairement dans les équations régissant l'état et l'observation en fonction de l'état, les coefficients de l'équation différentielle vérifiée par l'observation ne dépendent plus linéairement de ces paramètres. Ainsi ce modèle est un cas particulier de celui étudié dans la partie précédente. Nous assurons que des hypothèses classiques d'observabilité et de stabilité valident les conclusions de l'étude générale. Les résultats de cette partie ont été annoncés dans [33] A. LE BRETON, dans un contexte d'hypothèses particulières qui sont levées ici. Dans [7] A.V. BALAKRISHNAN a étudié ce même modèle sous l'hypothèse restrictive essentielle que la covariance du bruit d'observation est complètement connue dans le but d'utiliser la méthode du maximum de vraisemblance, sans pourtant donner de méthode d'estimation préalable des paramètres inconnus y figurant. Dans [6] A. BAGCHI étudie une méthode d'estimation simultanée de tous les paramètres basée sur la recherche d'un zéro du gradient d'une modification de la fonction de Log-vraisemblance construite à partir de la donnée d'une estimation a priori de la matrice de covariance du bruit d'observation. Dans ce travail,

nous levons la restriction majeure de l'étude de A.V. BALAKRISHNAN en donnant une méthode d'identification de cette covariance sur un intervalle de temps fini sans qu'il soit besoin, comme le suppose A. BAGCHI, d'en connaître une estimation a priori. En outre, le problème d'estimation des autres paramètres est alors relatif à un paramètre de dimension moindre que celle du paramètre mis en jeu dans l'approche de A. BAGCHI. Nous retrouvons en particulier les résultats de A.V. BALAKRISHNAN. De plus, nous assurons la normalité asymptotique de l'estimateur et donnons une caractérisation algébrique de la matrice intervenant dans la condition d'identifiabilité, rendant possible la vérification de la dite condition et permettant d'envisager une autre approche algorithmique de la construction de l'estimateur. Nous pensons donc, dans cette partie de notre travail avoir notablement amélioré les conclusions de l'étude faite dans [7] tant sur le plan théorique qu'en vue de leur utilisation pratique.

En guise de conclusion, nous discutons dans une courte troisième partie de quelques aspects du problème de la modélisation des systèmes stochastiques à temps continu et suggérons à cette occasion quelques prolongements possibles à notre travail.

CHAPITRE I

LE MODELE GENERAL - LE PROBLEME - LES OUTILS

A condition de ne pas se soucier de la précision du langage utilisé, on peut très bien présenter le thème de notre travail, en quelques lignes, de la manière suivante :

"On suppose que l'on observe au cours du temps une trajectoire d'un processus stochastique $(X_t ; t \geq 0)$ solution d'une équation différentielle stochastique de la forme :

$$\begin{cases} dX_t = a_\varphi(t, X)dt + \sigma_\varphi(t, X)dW_t & ; \quad t \geq 0 \\ X_0 = X(0) \end{cases}$$

où $(W_t ; t \geq 0)$ est un processus de Mouvement Brownien standard et $a_\varphi(\dots)$ et $\sigma_\varphi(\dots)$ désignent des applications dépendant d'un paramètre inconnu φ élément d'un ensemble $\Phi \subset \mathbb{R}^d$. On désire estimer la vraie valeur φ^* du paramètre φ au vu de cette observation, c'est-à-dire construire une famille d'applications définies sur l'espace des trajectoires possibles, à valeurs dans Φ , telle que, pour un mode de convergence de la théorie des probabilités, la famille d'éléments aléatoires ainsi obtenue converge vers la vraie valeur φ^* de φ ."

Dans la première partie de ce chapitre nous nous proposons de donner une formulation précise du problème dont nous venons de présenter grossièrement les éléments. En premier lieu, nous exposons ce que nous entendons par solution d'une équation différentielle stochastique ; la notion

de solution faible utilisée permet de construire directement la structure statistique associée au problème d'estimation de paramètres inconnus. Les théorèmes d'existence de solutions faibles d'équations différentielles stochastiques conduisent de façon naturelle à envisager, dans certains cas, l'étude du problème d'estimation par la méthode du maximum de vraisemblance ; la seconde partie du chapitre est alors consacrée à la présentation d'énoncés généraux concernant cette méthode, la troisième partie étant elle dévolue à l'étude de quelques propriétés des intégrales stochastiques préparant à l'application de ces énoncés dans le cadre du problème posé.

§ 1. - SOLUTIONS FAIBLE ET FORTE D'UNE EQUATION DIFFERENTIELLE STOCHASTIQUE - POSITION DU PROBLEME D'ESTIMATION.

1.1. Notations et définitions.

Etant donné p un entier, $p \geq 1$, et T un nombre réel, $T > 0$, on désigne par :

C l'espace des applications continues de \mathbb{R}^+ dans \mathbb{R}^p ,

S (resp. S_T) la tribu sur C engendrée par les applications coordonnées $\pi_t : x \rightsquigarrow x(t)$ pour $t \geq 0$ (resp. $0 \leq t \leq T$),

μ_* (resp. μ_*^T) la loi du mouvement brownien canonique ou mesure de Wiener standard sur $(C, (S_t)_{t \geq 0})$ (resp. $(C, (S_t)_{0 \leq t \leq T})$).

a et σ désignent des applications de $\mathbb{R}^+ \times C$ dans \mathbb{R}^p et de $\mathbb{R}^+ \times C$ dans $\mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$ respectivement. On suppose que les processus a et σ sont progressivement mesurables et adaptés à la famille $(S_t)_{t \geq 0}$ de tribus.

1.1.1. DEFINITION : "On appelle solution faible de l'équation différentielle stochastique

$$(1) \quad \begin{cases} dX_t = a(t, X)dt + \sigma(t, X)dW_t & ; \quad t \geq 0 \\ X_0 = z_0 \end{cases}$$

où $z_0 \in \mathbb{R}^p$ est donné, la donnée d'un couple $(P_{a, \sigma}, W^{a, \sigma})$ où $P_{a, \sigma}$ est une loi de probabilité sur (C, S) et $W^{a, \sigma} = (W_t^{a, \sigma}; t \geq 0)$ est un mouvement brownien standard sur $(C, (S_t)_{t \geq 0}, P_{a, \sigma})$ tel que, pour tout $t \geq 0$, on ait :

$$(2) \quad \Pi_t = z_0 + \int_0^t a(s, \cdot) ds + \int_0^t \sigma(s, \cdot) dW_s^{a, \sigma} \quad P_{a, \sigma} \text{ presque sûrement.}$$

On dit que la solution faible de (1) est unique si la loi $P_{a, \sigma}$ obtenue est unique."

Nous désignerons toujours par $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$ une base de processus où $(\mathcal{F}_t)_{t \geq 0}$ est une famille croissante de sous-tribus de la tribu \mathcal{F} sur Ω supposée P -complète, \mathcal{F}_0 contenant les ensembles de P -mesure nulle de \mathcal{F} . Par " $X = (X_t)_{t \geq 0}$ est un processus défini sur la base $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$ " nous signifions en particulier l'adaptation de X à la famille $(\mathcal{F}_t)_{t \geq 0}$.

1.1.2. DEFINITION : "Etant donné W un mouvement brownien standard défini sur une base $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$, on appelle solution forte de (1) un processus X défini sur cette même base tel que l'on ait :

$$(3) \quad X_t = z_0 + \int_0^t a(s, X) ds + \int_0^t \sigma(s, X) dW_s \quad P \text{ presque sûrement.}$$

On dit que la solution forte est unique si deux solutions X^1 et X^2 , processus sur $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$ sont indistinguables."

1.1.3. Remarques :

a) L'équation (1) n'est pas en général une équation différentielle stochastique de type "Ito", puisque $a(t, \cdot)$ et $\sigma(t, \cdot)$ peuvent dépendre du passé avant t des trajectoires. Il est bien connu, depuis la parution d'un article de YAMADA et WATANABE, que dans le cas d'équations du type "Ito", l'unicité de la solution forte (unicité trajectorielle) implique l'uni-

cit  de la solution faible (unicit  en loi) (voir par exemple P. PRIOURET [43]). La d monstration reste inchang e si on consid re des  quations du type (1) ; le r sultat de YAMADA-WATANABE reste donc vrai pour de telles  quations.

b) Si $(P_{a,\sigma}, W^{a,\sigma})$ est une solution faible de (1) et que $P_{a,\sigma}$ est la loi d'un processus X ,   trajectoires continues, d fini sur une base $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$, en posant :

$$\xi_t(\omega) = W_t^{a,\sigma}(X(\omega)) \quad , \quad \omega \in \Omega \quad ,$$

on obtient,   partir de (2), la relation :

$$X_t = z_0 + \int_0^t a(s, X) ds + \int_0^t \sigma(s, X) d\xi_s \quad P \text{ presque s urement.}$$

En d'autres termes, le processus X est solution forte de (1) relativement   la base $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, P)$ et au mouvement brownien $(\xi_t)_{t \geq 0}$.

c) Illustrons par un exemple tr s simple les approches "forte" et "faible" de la d finition de la solution d'une  quation diff rentielle stochastique. On se donne $(a, z_0) \in \mathbb{R}^p \times \mathbb{R}^p$ et $\sigma > 0$, et on consid re l' quation :

$$\begin{cases} dX_t = \sigma dW_t + a dt & ; \quad t \geq 0 \\ X_0 = z_0 \end{cases}$$

La solution faible est bien s r unique et consiste en le couple

$$(P_{a,\sigma}, (\frac{\Pi_t - z_0 - at}{\sigma})_{t \geq 0})$$

o , si $\underline{s} \in \mathcal{S}$,

$$P_{a,\sigma}(\underline{s}) = \mu_* \{x \in C : (t \rightsquigarrow z_0 + at + \sigma x_t) \in \underline{s}\} \quad .$$

L'approche "forte" consiste   se donner une base de processus $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F})$, et un mouvement brownien W sur cette base. La solution forte est unique et d finie par :

$$X_t = z_0 + at + \sigma W_t \quad ; \quad t \geq 0 \quad .$$

La loi de probabilité de cette solution forte coïncide évidemment avec $P_{a,\sigma}$. En outre le théorème de R.H. CAMERON et W.T. MARTIN (voir par exemple [15]), assure que les lois $P_{a,\sigma}/S_t$ et $P_{0,\sigma}/S_t$ sont équivalentes, la dérivée de Radon-Nikodym étant donnée par :

$$\frac{dP_{a,\sigma}}{dP_{0,\sigma}/S_t}(x) = \exp\left\{\frac{1}{\sigma^2} \langle a, x_t \rangle - \frac{1}{2\sigma^2} \|a\|_{t}^2\right\}.$$

Si $a \in \mathbb{R}^p$, $b \in \mathbb{R}^p$ on pose $\langle a, b \rangle = \sum_i a_i b_i$, $\|a\|^2 = \langle a, a \rangle$ et ${}^t a$ désigne le vecteur transposé de a . De même si $\sigma \in \mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$ on pose $\|\sigma\|^2 = \sum_{i,j} \sigma_{ij}^2 = \text{tr}(\sigma {}^t \sigma)$ où tr désigne la trace et ${}^t \sigma$ la transposée de σ .

Un théorème énoncé et démontré par I.V. GIRSANOV en 1960 (voir [24]) concernant la substitution de mesure dans une base de processus a été le point de départ de nombreux développements relatifs à l'équivalence et à l'absolue continuité des lois de processus de type diffusion. Le paragraphe suivant est consacré à la présentation des principaux résultats, obtenus par différents auteurs, que nous utiliserons dans la suite. Pour les démonstrations nous renvoyons aux auteurs. Pour un exposé systématique et détaillé on pourra consulter J. MEMIN [39].

1.2. Théorèmes d'équivalence et d'absolue continuité de lois sur (C, S_t) .

On rappelle le théorème suivant, maintenant classique :

1.2.1. THEOREME (I.V. GIRSANOV) : "Soit $(\Omega, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathcal{F}, P)$ une base de processus, \tilde{W} et W deux processus mesurables définis sur cette base, à valeurs dans \mathbb{R}^p , satisfaisant les propriétés suivantes :

(A1) W est un mouvement brownien p -dimensionnel ;

(A2) $\tilde{W}_t = \int_0^t A(s, \cdot) ds + W_t$ P presque sûrement ;

où (A3) A est une application de $[0, T] \times \Omega$ dans \mathbb{R}^p , mesurable, adaptée à la famille (\mathcal{F}_t) et telle que :

$$P\left[\int_0^T \|A(s, \cdot)\|^2 ds < +\infty\right] = 1 ;$$

$$(A4) \quad E[\zeta(\cdot)] = 1, \text{ où} \\ \zeta(\omega) = \exp\left\{-\int_0^T \langle A(s, \omega), dW_s(\omega) \rangle - \frac{1}{2} \int_0^T \|A(s, \omega)\|^2 ds\right\}.$$

Alors le processus \tilde{W} est un mouvement brownien défini sur la base $(\Omega, (\mathcal{F}_t), \tilde{P})$ où $\tilde{P} = \zeta \cdot P$."

Des généralisations de ce théorème ont été obtenues dans le cas des processus à valeurs hilbertiennes et dans le cas où le mouvement brownien est remplacé par une martingale continue de carré intégrable.

Nous nous intéressons essentiellement au cas où les hypothèses (A2) et (A3) sont remplacées par les suivantes :

$$(A2') \quad \tilde{W}_t = \int_0^t a(s, \tilde{W}) ds + W_t \quad P \text{ presque sûrement ;}$$

(A3') a est une application de $[0, T] \times C$ dans \mathbb{R}^p progressivement mesurable, adaptée à la famille (S_t) et telle que :

$$P\left[\int_0^T \|a(s, \tilde{W})\|^2 ds < +\infty\right] = 1.$$

L'énoncé suivant est un corollaire du théorème 1 de M.P. ERSHOV [19].

1.2.2. THEOREME : "Avec les hypothèses (A1), (A2'), (A3') et (A4), la loi $P_{\tilde{W}}$ du processus \tilde{W} est équivalente à la loi μ_*^T du mouvement brownien canonique, et dans l'espace (C, S_T) la dérivée de Radon-Nikodym est donnée par la relation :

$$(4) \quad \frac{dP_{\tilde{W}}}{d\mu_*^T}(x) = \exp\left[\int_0^T \langle a(s, x), d\tilde{W}_s(x) \rangle - \frac{1}{2} \int_0^T \|a(s, x)\|^2 ds\right] \\ \mu_*^T \text{ presque sûrement,}$$

où l'intégrale stochastique est prise par rapport à la base de processus $(C, (S_t)_{0 \leq t \leq T}, \mu_*^T)$."

Ce théorème admet la généralisation suivante (voir [39], Proposition 6.2.) :

1.2.3. THEOREME : "Soit Y et W des processus à trajectoires continues, à valeurs dans \mathbb{R}^p , définis sur une base $(\Omega, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathcal{F}, P)$, tels que :

(B1) W est un mouvement brownien ;

(B2) $P[\int_0^T (\|a(s, Y)\|^2 + \|\sigma(s, Y)\|^2) ds < +\infty] = 1$;

(B3) $\sigma(s, x)$ est une matrice inversible pour tout s et tout $x \in C$ et $\|\sigma(s, x)\| \geq k > 0$;

(B4) l'équation stochastique (5) $Z_t = \int_0^t \sigma(s, Z) dW_s$ admet une solution faible unique ;

(B5) $Y_t(\omega) = \int_0^t a(s, Y(\omega)) ds + \int_0^t \sigma(s, Y(\omega)) dW_s(\omega)$ P presque sûrement ;

(B6) $E[\zeta(\cdot)] = 1$, où
 $\zeta(\cdot) = \exp\{-\int_0^T \sigma^{-1}(s, Y) a(s, Y) dW_s\} - \frac{1}{2} \int_0^T \|\sigma^{-1}(s, Y) a(s, Y)\|^2 ds$

Alors la loi P_Y du processus Y est équivalente à la loi P_Z d'un processus solution de (5), la densité s'exprimant par la relation :

$$(6) \quad \frac{dP_Y}{dP_Z}(x) = \exp\left[\int_0^T \langle \sigma^{-1}(s, x) a(s, x), \sigma^{-1}(s, x) d\bar{W}_s(x) \rangle - \frac{1}{2} \int_0^T \|\sigma^{-1}(s, x) a(s, x)\|^2 ds\right]$$

P_Z -presque sûrement,

où l'intégrale stochastique est prise par rapport à la base de processus $(C, (\mathcal{S}_t)_{0 \leq t \leq T}, P_Z)$."

On donne alors des énoncés mettant en évidence les liens existant entre le problème d'équivalence de lois sur C et celui de l'existence et de l'unicité des solutions faibles de l'équation stochastique (1).

Les énoncés sont présentés pour des bases de processus de la forme $(C, (\mathcal{S}_t)_{t \in \mathbb{R}^+}, S, ..)$ et non $(C, (\mathcal{S}_t)_{0 \leq t \leq T}, S_T, ..)$ comme cela est fait par la plupart des auteurs cités. Toutefois ils sont des conséquences immédiates des énoncés originaux. En effet nous faisons les hypothèses convenables pour tout $T \geq 0$, aussi, si appliquant les dits énoncés pour

chaque $T \geq 0$, on construit une famille $\{\lambda_T; T \geq 0\}$ de mesures de probabilité définies respectivement sur (C, S_T) telle que, pour tout $t \in [0, T]$, λ_t soit la restriction de λ_T à S_t , le système $\{(C, S_T, \lambda_T); T \geq 0\}$ forme un système projectif d'espaces de probabilité (lorsqu'on prend pour application d'un espace dans l'autre l'application identique sur C) qui admet une limite projective λ sur (C, S) .

On commence par le cas particulier de :

$$(7) \quad X_t = z_0 + \int_0^t a(s, X) ds + W_t .$$

1.2.4. THEOREME (M.P. ERSHOV, [19] théorème 4) : "Supposons que :

$$(j) \quad \text{pour tout } x \in C \quad \text{on a} \quad \int_0^t \|a(s, x)\|^2 ds < +\infty \quad \text{pour tout } t > 0 .$$

Alors l'équation (7) admet une solution faible si et seulement si, on a :

$$(jj) \quad \text{pour tout } t > 0 , \\ E_{\mu_{*, z_0}} \left[\exp \left\{ \int_0^t \langle a(s, \cdot), d\Pi_s(\cdot) \rangle - \frac{1}{2} \int_0^t \|a(s, \cdot)\|^2 ds \right\} \right] = 1 , \\ \text{où, si } \underline{s} \in S , \quad \mu_{*, z_0}(\underline{s}) = \mu_* \{ x \in C : (t \rightsquigarrow z_0 + x_t) \in \underline{s} \} .$$

De plus, il y a alors unicité de la solution faible P_a et pour tout $t > 0$ les lois P_a/S_t et $\mu_{*, z_0}/S_t$ sont équivalentes, la dérivée de Radon-Nikodym étant donnée par :

$$(8) \quad \frac{dP_a}{d\mu_{*, z_0}} / S_t(x) = \exp \left\{ \int_0^t \langle a(s, x), d\Pi_s(x) \rangle - \frac{1}{2} \int_0^t \|a(s, x)\|^2 ds \right\} \\ \mu_{*, z_0} \text{ presque sûrement,}$$

où l'intégrale stochastique est prise relativement à la base de processus $(C, (S_t), S, \mu_{*, z_0})$."

Ce théorème se généralise comme suit dans le cas de l'équation (1) (voir [39] Théorème 6.5) :

1.2.5. THEOREME : "Soit a et σ avec les hypothèses (B3), (B4) et supposons que :

- (j) pour tout $x \in C$ on a $\int_0^t \|a(s,x)\|^2 ds < +\infty$ pour tout $t > 0$ et $Q_\sigma [\int_0^t \|\sigma(s,.)\|^2 ds < +\infty] = 1$ pour tout $t > 0$, Q_σ désignant la loi de la solution de
 (5)' $Z_t = z_0 + \int_0^t \sigma(s,Z) dW_s$.

Alors l'équation (1) admet une solution faible si et seulement si on a :

- (jj) pour tout $t > 0$,
 $E_{Q_\sigma} [\exp\{\int_0^t \langle \sigma^{-1}(s,.)a(s,.) , \sigma^{-1}(s,.) d\Pi_s \rangle - \frac{1}{2} \int_0^t \|\sigma^{-1}(s,.)a(s,.)\|^2 ds\}] = 1$.

De plus, il y a alors unicité de la solution faible $P_{a,\sigma}$ et pour tout $t > 0$ les lois $P_{a,\sigma}/S_t$ et Q_σ/S_t sont équivalentes, la densité étant donnée par :

$$(9) \quad \frac{dP_{a,\sigma}}{dQ_\sigma} / S_t (.) = \exp\{\int_0^t \langle \sigma^{-1}(s,.)a(s,.)\sigma^{-1}(s,.) , d\Pi_s \rangle - \frac{1}{2} \int_0^t \|\sigma^{-1}(s,.)a(s,.)\|^2 ds\}$$

Q_σ -presque sûrement,

où l'intégrale stochastique est prise relativement à la base de processus $(C, (S_t), S, Q_\sigma)$."

L'utilisation effective des théorèmes 1.2.4. et 1.2.5. nécessite que l'on soit capable de vérifier l'hypothèse (jj) des énoncés correspondants et en particulier, d'assurer l'existence d'une solution faible unique de l'équation (5)'. Dans cette perspective nous donnons, dans le paragraphe suivant des conditions suffisantes, facilement exploitables, d'existence d'une solution forte de l'équation (1) et de vérification de (jj).

1.3. Existences et unicités forte et faible des solutions de (1).

On indique d'abord un théorème d'existence et unicité forte pour les équations de type Ito (voir par exemple I.I. GIHMAN - A.V. SKOROHOD [23]).

1.3.1. THEOREME : "Soit $(\Omega, (\mathcal{F}_t), P)$ une base de processus
et W un mouvement brownien sur cette base, à valeurs dans \mathbb{R}^q .

Soit :

$$\begin{aligned} \tilde{\sigma} : \mathbb{R}^+ \times \mathbb{R}^p &\longrightarrow \mathfrak{L}(\mathbb{R}^q; \mathbb{R}^p) \quad \text{un champ de matrices } p \times q \text{ et} \\ \tilde{a} : \mathbb{R}^+ \times \mathbb{R}^p &\longrightarrow \mathbb{R}^p \quad \text{un champ de vecteurs} \end{aligned}$$

tels que :

- (i) $\forall t$, $\tilde{\sigma}$ et \tilde{a} en restriction à $[0, t] \times \mathbb{R}^p$ sont mesurables pour $\mathcal{B}_{[0, t]} \otimes \mathcal{B}_{\mathbb{R}^p}$
- (j) $\forall T > 0$, $\forall C > 0$, $\exists K_{T, C} > 0$ tel que :

$$\sup_{t \leq T} \{ \|\tilde{\sigma}(t, u) - \tilde{\sigma}(t, v)\| + \|\tilde{a}(t, u) - \tilde{a}(t, v)\| \} \leq K_{T, C} \|u - v\|$$
 $u, v \in \mathbb{R}^d$, $\|u\| \leq C$, $\|v\| \leq C$
- (k) $\forall T > 0$, $\exists K_T > 0$ tel que :

$$\sup_{t \leq T} \{ \|\tilde{\sigma}(t, u)\|^2 + \|\tilde{a}(t, u)\|^2 \} \leq K_T (1 + \|u\|^2)$$
 $u \in \mathbb{R}^d$.

Alors, pour tout $z_0 \in \mathbb{R}^p$, il existe un processus basé sur $(\Omega, (\mathcal{F}_t), P)$,
à valeurs dans \mathbb{R}^p , de norme au carré P-intégrable, défini sur $[0, +\infty[$,
à trajectoires continues, unique à l'indistingabilité près, tel que pour tout
 $t \geq 0$, la relation :

$$X_t = z_0 + \int_0^t \tilde{\sigma}(s, X_s) dW_s + \int_0^t \tilde{a}(s, X_s) ds$$

soit vraie P-presque sûrement."

Cet énoncé est applicable dans le cas de (1) lorsque :

$$\begin{aligned} a(t, x) &= \tilde{a}(t, x_t) \\ \sigma(t, x) &= \tilde{\sigma}(t, x_t) \end{aligned}$$

pour $x = (x_t ; t \geq 0) \in C$.

Il admet la généralisation suivante (voir par exemple P.S. LIPSER, A.N. SHIRYAYEV [36]).

1.3.2. THEOREME : "Soit $(\Omega, (\mathcal{F}_t), P)$ une base de processus
et W un mouvement brownien sur cette base. On suppose que :

$\forall T, \exists K_T > 0$ tel que : $\forall t \leq T$

$$(i) \quad \|a(t, x) - a(t, y)\|^2 + \|\sigma(t, x) - \sigma(t, y)\|^2 \\ \leq K_T \left\{ \int_0^t \|x_s - y_s\|^2 ds + \|x_t - y_t\|^2 \right\} \text{ et}$$

$$(j) \quad \|a(t, x)\|^2 + \|\sigma(t, x)\|^2 \leq K_T \left\{ \int_0^t (1 + \|x_s\|^2) ds + (1 + \|x_t\|^2) \right\} \text{ pour} \\ x, y \in C .$$

Alors, pour tout $z_0 \in \mathbb{R}^p$, il existe un processus basé sur $(\Omega, (\mathcal{F}_t), P)$,
à valeurs dans \mathbb{R}^p , de norme au carré P-intégrable, défini sur $[0, +\infty[$,
à trajectoires continues, unique à l'indistingabilité près, tel que, pour
tout $t \geq 0$, la relation (3) soit vraie P-presque sûrement."

Les deux théorèmes précédents fournissent, compte-tenu de la remarque 1.1.3.a., des hypothèses assurant l'existence et l'unicité de la solution faible de l'équation stochastique (1), donc en particulier de (5)'.

On s'intéresse maintenant à des conditions suffisantes pour que les hypothèses (jj) des théorèmes 1.2.4. et 1.2.5. soient vérifiées ; elles permettront de décrire un autre cadre d'hypothèses facilement exploitables dans lequel l'équation (1) admette une solution faible unique.

1.3.3. THEOREME : "Soit a et σ vérifiant les hypothèses
(B3) et (B4) et supposons que, pour tout $t > 0$,

$$Q_\sigma \left[\int_0^t \|\sigma(s, \cdot)\|^2 ds < +\infty \right] = 1 ,$$

Q_σ désignant la loi de la solution faible de (5).

L'une quelconque des trois conditions suivantes :

(C1) pour tout $T > 0$, a est borné uniformément en x sur
 $[0, T]$

(C2) pour tout $T > 0$, il existe $K_T > 0$ telle que :

$$\|a(t, x)\| \leq K_T (1 + |x|_T) \quad \text{pour tout } t \leq T, \quad x \in C \quad \text{où}$$

$$|x|_T = \sup_{0 \leq t \leq T} \|x_t\|$$

(C3) il existe $\delta > 0$ tel que :

$$\|a(t, x)\| \leq f_0(|x|_{t-\delta}) \quad \text{où } f_0 \text{ est une fonction croissante } \mathbb{R} \rightarrow \mathbb{R},$$

implique que, pour tout $t > 0$, on a :

$$E_{Q_\sigma} \left[\exp \left\{ \int_0^t \langle \sigma^{-1}(s, \cdot) a(s, \cdot), \sigma^{-1}(s, x) \rangle d\Pi_s \right\} - \frac{1}{2} \int_0^t \|\sigma^{-1}(s, \cdot) a(s, \cdot)\|^2 ds \right] = 1 . "$$

Cet énoncé fait la synthèse de résultats obtenus par I.V. GIRSANOV [24], V.E. BENES [10] et T. DUNCAN and P. VARAIYA [18].

Nous sommes alors en mesure de donner une formulation précise du problème statistique que nous envisageons dans la suite.

1.4. Position du problème d'estimation. Notions d'identifiabilité.

On se donne $z_0 \in \mathbb{R}^p$ et, Φ désignant un sous-ensemble pavé de l'espace euclidien \mathbb{R}^d , où d est un entier donné, $d \geq 1$, une famille $\{(a_\varphi, \sigma_\varphi) ; \varphi \in \Phi\}$ de couples d'applications telle que, pour tout $\varphi \in \Phi$, les hypothèses suivantes sont satisfaites :

$$(D1) \quad \begin{aligned} a_\varphi &: \mathbb{R}^+ \times C \rightarrow \mathbb{R}^p \\ \sigma_\varphi &: \mathbb{R}^+ \times C \rightarrow \mathcal{L}(\mathbb{R}^p; \mathbb{R}^p) \end{aligned}$$

sont des processus progressivement mesurables et adaptés à la famille $(S_t)_{t \geq 0}$ de tribus sur C .

(D2) Les équations différentielles stochastiques :

$$(1) \quad \begin{cases} dX_t = a_\varphi(t, X) dt + \sigma_\varphi(t, X) dW_t & ; \quad t \geq 0 \\ X_0 = z_0 \end{cases}$$

et (5)'

$$\begin{cases} dX_t = \sigma_\varphi(t, X) dW_t & ; \quad t \geq 0 \\ X_0 = z_0 \end{cases}$$

admettent chacune une solution faible unique notée respectivement P_φ et Q_φ . Pour tout $t > 0$ les lois P_φ/S_t et Q_φ/S_t sont équivalentes, la

dérivée de Radon-Nikodym étant donnée par :

$$(10) \frac{dP_\varphi}{dQ_\varphi} / S_t = \exp \left\{ \int_0^t \langle \sigma_\varphi^{-1}(s, \cdot) a_\varphi(s, \cdot), \sigma_\varphi^{-1}(s, \cdot) d\Pi_s \rangle - \frac{1}{2} \int_0^t \|\sigma_\varphi^{-1}(s, \cdot) a_\varphi(s, \cdot)\|^2 ds \right\}$$

Q_φ -presque sûrement, l'intégrale stochastique étant prise relativement à la base de processus $(C, (S_t)_{t \geq 0}, S, Q_\varphi)$.

On suppose que l'on observe continûment sur l'intervalle de temps $[0, T]$ où $T > 0$ est donné, une trajectoire $x = (x_t; t \geq 0)$ d'un processus stochastique dont, pour une valeur inconnue $\varphi^* \in \Phi$ dite vraie valeur du paramètre φ , P_φ^* soit la loi de probabilité. On se propose d'estimer φ^* au vu de cette observation.

Nous adoptons alors le formalisme de la statistique mathématique, tel qu'il a été introduit par J.R. BARRA dans son livre [9] pour formuler le problème statistique d'estimation du paramètre inconnu φ . La structure statistique du problème est :

$$(C, S_T, \{P_\varphi^T; \varphi \in \Phi\})$$

où, pour tout $\varphi \in \Phi$, $P_\varphi^T = P_\varphi / S_T$. Il s'agit alors de construire des estimateurs de φ i.e. des statistiques, définies sur cette structure, à valeurs dans Φ , possédant des propriétés souhaitables, en particulier de convergence. On introduit diverses notions d'identifiabilité. Si $\Delta = \{1, 2, \dots, r\}$, où $1 \leq r \leq d$, on note, pour $\varphi = (\varphi_1, \dots, \varphi_d) \in \mathbb{R}^d$:

$$\varphi_\Delta = (\varphi_1, \dots, \varphi_r) \quad \text{et} \quad \varphi_{\bar{\Delta}} = (\varphi_{r+1}, \dots, \varphi_d).$$

1.4.1. DEFINITION : "On dit que, dans la paramétrisation définie sur Φ , $\varphi \rightsquigarrow (a_\varphi, \sigma_\varphi)$, la ("vraie") valeur φ_Δ^* est fortement identifiable presque sûrement (resp. en moyenne quadratique, resp. en probabilité) sur l'intervalle de temps $[0, T]$ s'il existe une suite $\{\hat{\varphi}_{\Delta, T}^N; N \geq 1\}$ de statistiques définies sur la structure $(C, S_T, \{P_\varphi^T; \varphi \in \Phi\})$ telle que, pour tout $\varphi = (\varphi_\Delta, \varphi_{\bar{\Delta}}) \in \Phi$ où $\varphi_\Delta = \varphi_\Delta^*$, on ait :

$$\lim_{N \rightarrow +\infty} \hat{\varphi}_{\Delta, T}^N = \varphi_{\Delta}^*$$

au sens de la convergence P_{φ}^T -presque sûre (resp. en moyenne quadratique, resp. en probabilité, relativement à la mesure P_{φ}^T). On dit alors que le paramètre φ_{Δ}^* est identifié sur $[0, T]$ par la suite $\{\hat{\varphi}_{\Delta, T}^N; N \geq 1\}$ de statistiques."

1.4.2. DEFINITION : "On dit que, dans la paramétrisation, définie sur Φ , $\varphi \rightsquigarrow (a_{\varphi}, \sigma_{\varphi})$, la ("vraie") valeur φ^* est asymptotiquement identifiable presque sûrement (resp. en moyenne quadratique, resp. en probabilité) s'il existe une famille $\{\hat{\varphi}_T; T > 0\}$ de statistiques définies respectivement sur les structures $(C, S_T, \{P_{\varphi}; \varphi \in \Phi\})$, $T > 0$, telle que l'on ait :

$$\lim_{T \rightarrow +\infty} \tilde{\varphi}_T = \varphi^*$$

au sens de la convergence P_{φ^*} -presque sûre (resp. en moyenne quadratique, resp. en probabilité, relativement à la mesure P_{φ^*})."

Si dans la paramétrisation $\varphi \rightsquigarrow (a_{\varphi}, \sigma_{\varphi})$ la valeur φ_{Δ}^* est fortement identifiable presque sûrement sur $[0, T]$, il est légitime, l'ayant évaluée (avec probabilité 1) par $\lim_{N \rightarrow +\infty} \hat{\varphi}_{\Delta, T}^N$, de la supposer connue et d'étudier le problème d'estimation (ou d'identification) du seul paramètre vectoriel inconnu φ_{Δ}^* de dimension $d-r$.

Ainsi si $\sigma_{\varphi} \equiv \sigma$ ne dépend pas de paramètres inconnus (éventuellement parce que les paramètres y figurant auront été identifiés) la structure statistique $(C, S_T, \{P_{\varphi}^T; \varphi \in \Phi\})$ est dominée par la mesure Q_{σ}^T restriction à S_T de la mesure Q_{σ} , solution faible unique de l'équation :

$$(11) \quad \begin{cases} dX_t = \sigma(t, X) dW_t & ; \quad t \geq 0 \\ X_0 = z_0 \end{cases}$$

avec pour fonction de vraisemblance :

$$(12) \quad \frac{dP_{\varphi}^T}{dQ_{\sigma}^T} = \exp\left\{\int_0^T \langle \sigma^{-1}(t, \cdot) a_{\varphi}(t, \cdot), \sigma^{-1}(t, \cdot) d\Pi_t \rangle - \frac{1}{2} \int_0^T \|\sigma^{-1}(t, \cdot) a_{\varphi}(t, \cdot)\|^2 dt\right\} .$$

Il est alors naturel d'aborder le problème d'estimation de φ par la méthode du maximum de vraisemblance, i.e. de rechercher, si tant est qu'elle existe, une statistique $\tilde{\varphi}_T$ telle que :

$$(13) \quad \frac{dP_{\tilde{\varphi}_T}^T}{dQ_{\sigma}^T}(x) = \sup_{\varphi \in \Phi} \frac{dP_{\varphi}^T}{dQ_{\sigma}^T}(x) \quad , \quad x \in C \quad ,$$

ou peut-être

$$(13)' \quad \frac{d}{d\varphi} \left(\log \frac{dP_{\varphi}^T}{dQ_{\sigma}^T}(\tilde{\varphi}_T(x)) \right) = 0 \quad , \quad x \in C \quad .$$

De la même façon, lorsque l'observation a lieu en temps discret, la méthode du maximum de vraisemblance paraît adaptée dans ce cas.

La seconde partie de ce chapitre est alors consacrée à la présentation d'énoncés généraux sur l'estimation par maximum de vraisemblance dans une structure statistique où l'observation porte sur une fonction aléatoire en temps continu, que cette observation ait lieu en temps continu ou en temps discret.

§ 2. - ESTIMATION PAR LA METHODE DU MAXIMUM DE VRAISEMBLANCE LORSQUE L'OBSERVATION PORTE SUR UNE FONCTION ALEATOIRE EN TEMPS CONTINU.

2.1. Quelques lemmes.

De nombreux auteurs se sont intéressés à l'existence et à la construction de suites convergentes d'estimateurs du maximum de vraisemblance dans des structures statistiques où l'observation porte sur une suite d'éléments aléatoires indépendants ou non. On peut consulter, par exemple, J. AITCHINSON et S.D. SILVEY [1] et P. BILLINGSLEY [11].

Que l'estimation soit basée sur la fonction de vraisemblance exacte ou une approximation de celle-ci, les résultats obtenus reposent essentiellement sur l'utilisation d'un énoncé figurant dans [1] qui nous sera également utile ici.

2.1.1. LEMME [1] : "Soit ℓ une application de \mathbb{R}^d dans \mathbb{R}^d , continue sur la boule fermée $B(\rho) = \{u \in \mathbb{R}^d : \|u\| \leq \rho\}$, $\rho > 0$, vérifiant la condition :

$$(i) \quad \forall u \in \mathbb{R}^d, \quad \|u\| = \rho : \langle \ell(u), u \rangle < 0 .$$

Alors, il existe au moins une solution u^* de l'équation :

$$\ell(u) = 0$$

qui satisfasse $\|u^*\| < \rho$."

Démonstration : Supposons que :

$$\forall u \in B(\rho) : \ell(u) \neq 0 .$$

On peut alors définir sur $B(\rho)$ l'application λ par :

$$\lambda(u) = \rho \cdot \frac{\ell(u)}{\|\ell(u)\|}$$

λ est une application continue de $B(\rho)$ dans $B(\rho)$. Le théorème de BROUWER (voir par exemple J. WARGA [52] p.163) assure alors qu'il existe au moins une solution $\tilde{u} \in B(\rho)$ de l'équation :

$$\lambda(u) = u .$$

On a de plus $\|\tilde{u}\| = \|\lambda(\tilde{u})\| = \rho$.

On évalue maintenant :

$$\begin{aligned} \langle \ell(\tilde{u}), \tilde{u} \rangle &= \left\langle \frac{1}{\rho} \|\ell(\tilde{u})\| \cdot \lambda(\tilde{u}), \tilde{u} \right\rangle \\ &= \frac{\|\ell(\tilde{u})\|}{\rho} \cdot \|\tilde{u}\|^2 \geq 0 . \end{aligned}$$

Ce qui est en contradiction avec la condition (i). Ainsi il existe au moins une solution u^* de l'équation :

$$\ell(u) = 0$$

qui satisfasse $\|u^*\| \leq \rho$.

Alors $\langle \ell(u^*), u^* \rangle = 0$;

donc $\|u^*\| < \rho$. Ceci achève la démonstration. ■

Nous démontrons maintenant un lemme de mesurabilité qui nous permettra d'assurer que la résolution des équations (13) et (13)' fournit bien des statistiques.

2.1.2. LEMME : "Soit (Ω, \mathcal{A}) un espace mesurable. Soit ℓ une application de $\Omega \times \mathbb{R}^d$ dans \mathbb{R}^d telle que pour $w \in \Omega$, w fixé, $\ell(w, \cdot)$ est continue et pour $u \in \mathbb{R}^d$, u fixé, $\ell(\cdot, u)$ est mesurable. Soit h une application mesurable de Ω dans \mathbb{R}^d . On suppose que pour tout $w \in \Omega$ il existe $r(w) \geq 0$ tel que dans la boule $B(h(w), r(w)) = \{u \in \mathbb{R}^d : \|h(w) - u\| \leq r(w)\}$ l'équation :

$$\ell(w, u) = 0$$

admet une solution unique $u(w)$ (i.e. $u(w)$ est l'unique solution de $\ell(w, u) = 0$ qui soit la plus proche de $h(w)$).

Alors l'application $u : w \rightsquigarrow u(w)$ est \mathcal{A} -mesurable."

Démonstration : En vertu des hypothèses faites sur l'application ℓ , l'ensemble $\{(w, u) \in \Omega \times \mathbb{R}^d : \ell(w, u) = 0\}$ est $\mathcal{A} \otimes \mathcal{B}_{\mathbb{R}^d}$ -mesurable. Ainsi le graphe de la multiapplication A définie par

$$A(w) = \{u \in \mathbb{R}^d : \ell(w, u) = 0\} , \quad w \in \Omega ,$$

qui est à valeurs fermées en raison de la continuité de $\ell(w, \cdot)$, est mesurable. A est donc mesurable et, par suite, admet une suite

$\{u_n ; n \in \mathbb{N}\}$ dense de sections mesurables i.e. $\forall w \in \Omega$,

$A(w) = \overline{\{u_n(w) ; n \in \mathbb{N}\}}$, où pour tout $n \in \mathbb{N}$, $w \rightsquigarrow u_n(w)$ est mesurable.

On considère alors :

$$d(h(w), A(w)) = \inf\{\|h(w) - u\| ; u \in A(w)\} .$$

D'après ce qui précède, on a :

$$d(h(w), A(w)) = \inf\{\|h(w) - u_n(w)\| ; n \in \mathbb{N}\} .$$

L'application $w \rightsquigarrow d(h(w), A(w))$ est donc mesurable. Alors l'ensemble $\{(w, u) \in \Omega \times \mathbb{R}^d : d(h(w), A(w)) = \|h(w) - u\|\}$ est $\mathcal{A} \otimes \mathcal{B}_{\mathbb{R}^d}$ -mesurable et la multiapplication Γ , définie par :

$$\Gamma(w) = \{u \in \mathbb{R}^d : d(h(w), A(w)) = \|h(w) - u\|\} , \quad w \in \Omega ,$$

est mesurable. Par suite, la multiapplication $w \rightsquigarrow \Gamma(w) \cap A(w) = \{u(w)\}$ est mesurable et l'application $w \rightsquigarrow u(w)$ est mesurable. ■

2.2. Cas de l'observation en temps continu.

On considère une structure statistique asymptotique :

$$(\Omega, \mathcal{A}, (\mathcal{A}_T)_{T \in \mathbb{R}^+}, \{P_\varphi ; \varphi \in \Phi\})$$

où (Ω, \mathcal{A}) est un espace probablisable dit espace des observations ;

$(\mathcal{A}_T)_{T \geq 0}$ est une famille croissante de sous-tribus de \mathcal{A} où, pour tout $T \geq 0$, \mathcal{A}_T représente la tribu des événements observables jusqu'à l'instant T ;

$\{P_\varphi ; \varphi \in \Phi\}$ désigne la famille des lois de probabilité possibles régissant les observations, indexée par un paramètre φ parcourant un sous-ensemble Φ de l'espace euclidien \mathbb{R}^d .

On suppose qu'il existe une loi de probabilité ν sur (Ω, \mathcal{A}) telle que pour tout $T > 0$, la mesure ν^T , restriction de ν à \mathcal{A}_T , domine la famille $\{P_\varphi^T ; \varphi \in \Phi\}$ des restrictions à \mathcal{A}_T des lois P_φ .

On appelle fonction de Log-vraisemblance une version $L_T(\varphi, .)$ de

$$\text{Log} \frac{dP_\varphi^T}{d\nu^T}$$

telle que, pour tout $w \in \Omega$, w fixé, l'application $L_T(., w)$ est deux

fois continûment différentiable sur Φ . On suppose qu'une telle version existe.

On désignera respectivement par $L_T^{(1)}(\varphi, \cdot)$ et $L_T^{(2)}(\varphi, \cdot)$ le vecteur des dérivées partielles premières $\frac{\partial}{\partial \varphi_i} L_T(\varphi, \cdot)$ $i=\overline{1, d}$ et la matrice des dérivées partielles secondes $\frac{\partial^2}{\partial \varphi_j \partial \varphi_i} L_T(\varphi, \cdot)$, $i, j=\overline{1, d}$.

On énonce alors une proposition qui généralise celle annoncée dans A. LE BRETON [32] à propos d'un modèle particulier :

2.2.1. PROPOSITION. " $\varphi^* \in \Phi$ désignant la vraie valeur du paramètre φ , on fait les hypothèses suivantes :

- (E1) $P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \frac{1}{T} L_T^{(1)}(\varphi^*, \cdot) = 0$,
- (E2) $P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \frac{1}{T} L_T^{(2)}(\varphi^*, \cdot) = \Gamma$, où Γ est une matrice définie négative,
- (E3) $\mathcal{L}_{P_{\varphi^*}} \text{-} \lim_{T \rightarrow +\infty} \frac{1}{T^{1/2}} L_T^{(1)}(\varphi^*, \cdot) = N(0, -\Gamma)$, où $N(0, \Lambda)$ désigne un vecteur aléatoire de la loi gaussienne centrée, de matrice de covariances Λ ,
- (E4) Pour tout $T > 0$ et $\eta > 0$ il existe une variable aléatoire $\varepsilon_{T, \eta}$ positive finie \mathcal{C}_T -mesurable, telle que :
- $$\sup_{\varphi: \|\varphi - \varphi^*\| \leq \eta} \frac{1}{T} \|L_T^{(2)}(\varphi, \cdot) - L_T^{(2)}(\varphi^*, \cdot)\| \leq \varepsilon_{T, \eta}.$$

De plus $\forall \rho > 0 \quad \lim_{\eta \rightarrow 0} \limsup_{T \rightarrow +\infty} P_{\varphi^*}[\varepsilon_{T, \eta} > \rho] = 0$.

Alors il existe une famille $\{\tilde{\varphi}_T; T > 0\}$ d'estimateurs de φ vérifiant :

- (i) pour tout $T > 0$, $\tilde{\varphi}_T$ est \mathcal{C}_T -mesurable ;
- (ii) $\lim_{T \rightarrow +\infty} P_{\varphi^*}[L_T^{(1)}(\tilde{\varphi}_T, \cdot) = 0] = 1$;
- (iii) $P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \tilde{\varphi}_T = \varphi^*$ i.e. φ^* est asymptotiquement identifiable en probabilité par $\{\tilde{\varphi}_T; T > 0\}$.

En outre, si $\{\tilde{\varphi}'_T ; T > 0\}$ désigne une autre famille d'estimateurs vérifiant les conditions (i), (ii) et (iii), alors :

$$(iv) \quad \lim_{T \rightarrow +\infty} P_{\varphi^*} [\tilde{\varphi}'_T = \tilde{\varphi}_T] = 1 .$$

Enfin, si $\{\tilde{\varphi}_T ; T > 0\}$ est une famille d'estimateurs vérifiant (i), (ii) et (iii), on a :

$$(v) \quad \mathfrak{L}_{P_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{1/2} (\tilde{\varphi}_T - \varphi^*) = N(0, -\Gamma^{-1})$$

et (vi) $\mathfrak{L}_{P_{\varphi^*}} - \lim_{T \rightarrow +\infty} 2[L_T(\tilde{\varphi}_T, \cdot) - L_T(\varphi^*, \cdot)] = \chi_d^2$, où χ_d^2 désigne une variable aléatoire de loi du Khi deux à d degrés de liberté.

Dans cet énoncé et dans toute la suite :

P-lim désigne "limite au sens de la convergence en probabilité P " ;

\mathfrak{L}_P -lim désigne "limite au sens de la convergence en loi relativement à P " .

La démonstration de ce théorème s'appuiera sur deux lemmes.

2.2.2. LEMME : "Il existe une famille $\{\tilde{\epsilon}_T(\varphi) ; T > 0, \varphi \in \Phi\}$ de variables aléatoires telle que pour tout $T > 0$ et $\eta > 0$:

$$\sup_{\varphi: \|\varphi - \varphi^*\| \leq \eta} \|\tilde{\epsilon}_T(\varphi)\| \leq \tilde{\epsilon}_{T, \eta}$$

où $\{\tilde{\epsilon}_{T, \eta} ; T > 0, \eta > 0\}$ est une famille de variables aléatoires vérifiant :

(j) $\forall T > 0$, $\tilde{\epsilon}_{T, \eta}$ est \mathcal{O}_T -mesurable ;

(jj) $\forall \rho > 0$, $\lim_{\eta \rightarrow 0} \limsup_{T \rightarrow +\infty} P_{\varphi^*} [\tilde{\epsilon}_{T, \eta} > \rho] = 0$;

et telle que l'on ait, pour $\varphi \in \Phi$,

$$\frac{1}{T} L_T^{(1)}(\varphi, \cdot) = \frac{1}{T} L_T^{(1)}(\varphi^*, \cdot) + \Gamma(\varphi - \varphi^*) + \tilde{\epsilon}_T(\varphi)(\varphi - \varphi^*) . "$$

Démonstration : On écrit, pour $\varphi \in \Phi$,

$$\frac{1}{T}L_T^{(1)}(\varphi, \cdot) = \frac{1}{T}L_T^{(1)}(\varphi^*, \cdot) + \frac{1}{T}L_T^{(2)}(\varphi^*, \cdot)(\varphi - \varphi^*) + \left\{ \frac{1}{T}L_T^{(2)}(\varphi'(\varphi, \cdot), \cdot) - \frac{1}{T}L_T^{(2)}(\varphi^*, \cdot) \right\}(\varphi - \varphi^*)$$

où $\varphi'(\varphi, \cdot) = \varphi^* + \alpha(\varphi, \cdot)(\varphi - \varphi^*)$; $\alpha(\varphi, \cdot) \in [0, 1]$.

$$\begin{aligned} \text{Alors } \frac{1}{T}L_T^{(1)}(\varphi, \cdot) &= \frac{1}{T}L_T^{(1)}(\varphi^*, \cdot) + \Gamma(\varphi - \varphi^*) + \left\{ \frac{1}{T}L_T^{(2)}(\varphi^*, \cdot) - \Gamma \right\}(\varphi - \varphi^*) \\ &\quad + \left\{ \frac{1}{T}L_T^{(2)}(\varphi'(\varphi, \cdot), \cdot) - \frac{1}{T}L_T^{(2)}(\varphi^*, \cdot) \right\}(\varphi - \varphi^*) \end{aligned}$$

et posant

$$\tilde{\varepsilon}_T(\varphi) = \left[\frac{1}{2}L_T^{(2)}(\varphi^*, \cdot) - \Gamma \right] + \left[\frac{1}{T}L_T^{(2)}(\varphi'(\varphi, \cdot), \cdot) - \frac{1}{2}L_T^{(2)}(\varphi^*, \cdot) \right]$$

on a, compte-tenu de l'hypothèse (E4) :

$$\sup_{\varphi: \|\varphi - \varphi^*\| \leq \eta} \|\tilde{\varepsilon}_T(\varphi)\| \leq \left\| \frac{1}{T}L_T^{(2)}(\varphi^*, \cdot) - \Gamma \right\| + \varepsilon_{T, \eta} = \tilde{\varepsilon}_{T, \eta}$$

où, en vertu de (E2) et (E4),

$$\forall \rho > 0, \quad \lim_{\eta \rightarrow 0} \limsup_{T \rightarrow +\infty} P_{\varphi^*}[\varepsilon_{T, \eta} > \rho] = 0 \quad \blacksquare$$

2.2.3. LEMME : "Il existe une famille $\{\Omega(\eta, T) ; \eta > 0, T > 0\}$ d'événements vérifiant :

- (i) $\Omega(\eta, T) \in \mathcal{A}_T$
- (ii) $\lim_{\eta \rightarrow 0} \liminf_{T \rightarrow +\infty} P_{\varphi^*}[\Omega(\eta, T)] = 1$

telle que, pour tout $\eta > 0$, tout $T > 0$ et tout $\omega \in \Omega(\eta, T)$ l'équation

$$L_T^{(1)}(\varphi, \omega) = 0$$

admet une solution unique $\hat{\varphi}_T(\omega)$ dans la boule $\|\varphi - \varphi^*\| \leq \eta$, cette solution vérifiant de plus

$$L_T(\hat{\varphi}_T(\omega), \omega) > L_T(\varphi, \omega)$$

pour tout $\varphi \neq \hat{\varphi}_T(\omega)$ tel que $\|\varphi - \varphi^*\| \leq \eta$."

Démonstration : D'après le lemme précédent, si on pose $\varphi - \varphi^* = u$, on peut écrire :

$$\langle \frac{1}{T} L_T^{(1)}(\varphi^* + u, \omega), u \rangle = \langle \frac{1}{T} L_T^{(1)}(\varphi^*, \omega), u \rangle + {}^t u \Gamma u + {}^t u \tilde{\varepsilon}_T(\varphi^* + u, \omega) u .$$

Comme, par hypothèse (E2), $(-\Gamma)$ est définie positive, on a :

$$\inf_{u: \|u\|=1} \{-{}^t u \Gamma u\} = \gamma > 0$$

et
$$-{}^t u \Gamma u \geq \gamma \|u\|^2, \quad u \in \mathbb{R}^d .$$

Ainsi $\langle \frac{1}{T} L_T^{(1)}(\varphi^* + u, \omega), u \rangle$ sera strictement négatif pour $\|u\| = \eta$ si on a simultanément :

$$\|\frac{1}{T} L_T^{(1)}(\varphi^*, \omega)\| < \frac{\gamma \eta}{2} \quad \text{et} \quad \sup_{\|u\| \leq \eta} \|\tilde{\varepsilon}_T(\varphi^* + u, \omega)\| < \frac{\gamma}{2} .$$

On considère alors, pour $\eta > 0$, $T > 0$:

$$\Omega(\eta, T) = \{\omega \in \Omega : \|\frac{1}{T} L_T^{(1)}(\varphi^*, \omega)\| < \frac{\gamma \eta}{2} \text{ et } \tilde{\varepsilon}_{T, \eta}(\omega) < \frac{\gamma}{2}\} .$$

On a bien sûr :

$$\Omega(\eta, T) \in \mathfrak{a}_T \quad \text{et, pour } \omega \in \Omega(\eta, T)$$

$$\langle \frac{1}{T} L_T^{(1)}(\varphi^* + u, \omega), u \rangle < 0 \quad \text{pour } \|u\| = \eta .$$

D'autre part l'application $u \rightsquigarrow \frac{1}{T} L_T^{(1)}(\varphi^* + u, \omega)$ est continue sur $\|u\| \leq \eta$.

Le lemme 2.1.1. assure donc que l'équation :

$$\frac{1}{T} L_T^{(1)}(\varphi^* + u, \omega) = 0$$

admet une solution $U_T(\omega)$ vérifiant $\|U_T(\omega)\| < \eta$.

D'autre part, pour $\omega \in \Omega(\eta, T)$, on a aussi pour $\|u\| \leq \eta$,

$$\frac{1}{T} L_T^{(2)}(\varphi^* + u, \omega) - \Gamma = [\frac{1}{T} L_T^{(2)}(\varphi^* + u, \omega) - \frac{1}{T} L_T^{(2)}(\varphi^*, \omega)] + [\frac{1}{T} L_T^{(2)}(\varphi^*, \omega) - \Gamma]$$

puis
$$\|\frac{1}{T} L_T^{(2)}(\varphi^* + u, \omega) - \Gamma\| \leq \tilde{\varepsilon}_{T, \eta} < \frac{\gamma}{2}$$

et par conséquent $L_T^{(2)}(\varphi^* + u, \omega)$ est définie négative.

Calculant, pour $\omega \in \Omega(\eta, T)$ et $\|u\| \leq \eta$,

$$L_T^{(1)}(\varphi^* + u, \omega) - L_T^{(1)}(\varphi^* + U_T(\omega), \omega) = \int_0^1 \frac{d}{d\alpha} [L_T^{(1)}(\varphi^* + U_T(\omega) + \alpha(u - U_T(\omega)), \omega)] d\alpha$$

soit :

$$L_T^{(1)}(\varphi^* + u, \omega) = \int_0^1 L_T^{(2)}(\varphi^* + U_T(\omega) + \alpha(u - U_T(\omega)), \omega)(u - U_T(\omega)) d\alpha$$

il vient :

$$\begin{aligned} \langle L_T^{(1)}(\varphi^* + u, \omega), u - U_T(\omega) \rangle &= \int_0^1 t(u - U_T(\omega)) L_T^{(2)}(\varphi^* + U_T(\omega) + \alpha(u - U_T(\omega)), \omega)(u - U_T(\omega)) d\alpha \\ &< 0 \quad \text{si } u \neq U_T(\omega) . \end{aligned}$$

Ainsi $L_T^{(1)}(\varphi^* + u, \omega) \neq 0$ pour $\|u\| \leq \eta$ et $u \neq U_T(\omega)$.

D'autre part, $\hat{L}_T(\varphi^* + \cdot, \omega)$ atteint évidemment son maximum sur $\|u\| \leq \eta$ en $U_T(\omega)$.

Le résultat du lemme s'acquiert en posant $\hat{\varphi}_T(\omega) = \varphi^* + U_T(\omega)$ sous réserve de vérification de (ii).

Or, considérons l'inégalité :

$$P_{\varphi^*}[\Omega(\eta, T)] \leq P_{\varphi^*}[\|\frac{1}{T} L_T^{(1)}(\varphi^*, \cdot)\| \geq \frac{\gamma\eta}{2}] + P_{\varphi^*}[\tilde{\varepsilon}_{T, \eta}(\cdot) \geq \frac{\gamma}{2}] .$$

En vertu de l'hypothèse (E1) on a :

$$\limsup_{T \rightarrow +\infty} P_{\varphi^*}[\Omega(\eta, T)] \leq \limsup_{T \rightarrow +\infty} P_{\varphi^*}[\tilde{\varepsilon}_{T, \eta}(\cdot) \geq \frac{\gamma}{2}]$$

et, compte-tenu de (jj) lemme 2.2.2., il vient :

$$\lim_{\eta \rightarrow 0} \limsup_{T \rightarrow +\infty} P_{\varphi^*}[\Omega(\eta, T)] = 0$$

ou

$$\lim_{\eta \rightarrow 0} \liminf_{T \rightarrow +\infty} P_{\varphi^*}[\Omega(\eta, T)] = 1 \quad \blacksquare$$

2.2.4. Démonstration de la proposition 2.2.1.

a) Points (i) jusqu'à (iv).

Posons $\Omega_T = \bigcup_{n>0} \Omega(\frac{1}{n}, T) \in \mathcal{A}_T$ et considérons $\omega \in \Omega_T$. D'après

le lemme 2.2.3. il existe un entier $n(w) > 0$ tel que l'équation $L_T^{(1)}(w, x) = 0$ ait une solution unique $\hat{\varphi}_T(w)$ dans la boule $\|\varphi - \varphi^*\| \leq \frac{1}{n(w)}$. Nous sommes donc dans les conditions d'application du lemme 2.1.2. en y faisant :

$$\begin{aligned}(\Omega, \mathcal{A}) &= (\Omega_T, \mathcal{A}_T \cap \Omega_T) \\ \ell(w, u) &= L_T^{(1)}(u, w) \\ \text{et } h(w) &\equiv \varphi^*\end{aligned}$$

et, par conséquent, l'application $w \rightsquigarrow \hat{\varphi}_T(w)$ définie sur $(\Omega_T, \mathcal{A}_T \cap \Omega_T)$ est $\mathcal{A}_T \cap \Omega_T$ -mesurable.

On pose alors :

$$\tilde{\varphi}_T(w) = \begin{cases} \hat{\varphi}_T(w) & \text{si } w \in \Omega_T \\ \varphi_0 & \text{si } w \in \complement \Omega_T \end{cases} .$$

où φ_0 est une valeur quelconque du paramètre.

Il est clair que l'application $w \rightsquigarrow \tilde{\varphi}_T(w)$ est \mathcal{A}_T -mesurable.

On évalue maintenant pour $\rho > 0$ donné :

$$\lim_{T \rightarrow +\infty} P_{\varphi^*}[\|\tilde{\varphi}_T - \varphi^*\| \geq \rho] .$$

Comme pour tout $n \in \mathbb{N}^*$ tel que $\frac{1}{n} < \rho$ on a :

$$\Omega\left(\frac{1}{n}, T\right) \subset [\|\tilde{\varphi}_T - \varphi^*\| < \rho]$$

on a aussi $P_{\varphi^*}[\|\tilde{\varphi}_T - \varphi^*\| < \rho] \geq P_{\varphi^*}[\Omega\left(\frac{1}{n}, T\right)]$.

Or, pour tout $\varepsilon > 0$, $\exists \eta_0 > 0$ tel que :

$$\eta < \eta_0 \Rightarrow \liminf_{T \rightarrow +\infty} P_{\varphi^*}[\Omega(\eta, T)] > 1 - \varepsilon .$$

Choisissant alors n tel que $\frac{1}{n} < \rho$ et $\frac{1}{n} < \eta_0$, il existe T_0 tel que

$$T > T_0 \Rightarrow P_{\varphi^*}[\Omega\left(\frac{1}{n}, T\right)] > 1 - \varepsilon$$

puis

$$P_{\varphi^*}[\|\tilde{\varphi}_T - \varphi^*\| < \rho] > 1 - \varepsilon .$$

Ainsi $P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \tilde{\varphi}_T = \varphi^*$ et le point (iii) est démontré.

D'autre part, pour tout $n \geq 1$, on a :

$$\Omega\left(\frac{1}{n}, T\right) \subset [L_T^{(1)}(\hat{\varphi}_T(\cdot), \cdot) = 0]$$

donc $P_{\varphi^*}[L_T^{(1)}(\hat{\varphi}_T(\cdot), \cdot) = 0] \geq P_{\varphi^*}[\Omega\left(\frac{1}{n}, T\right)] \geq 1 - \epsilon$

dès que $T > T_0$ où T_0 est celui choisi plus haut.

Ainsi $\lim_{T \rightarrow +\infty} P_{\varphi^*}[L_T^{(1)}(\hat{\varphi}_T(\cdot), \cdot) = 0] = 1$ et (ii) est acquis.

Enfin, considérons une autre famille d'estimateurs $\{\tilde{\varphi}'_T; T > 0\}$ vérifiant (i) jusqu'à (iii).

On a, pour tout $n \geq 1$,

$$\{\omega \in \Omega : \omega \in \Omega\left(\frac{1}{n}, T\right), \|\varphi'_T(\omega) - \varphi^*\| \leq \frac{1}{n}, L_T^{(1)}(\varphi'_T(\omega), \omega) = 0\} \subset [\tilde{\varphi}'_T = \tilde{\varphi}_T]$$

donc $P_{\varphi^*}[\tilde{\varphi}'_T \neq \tilde{\varphi}_T] \leq (1 - P_{\varphi^*}[\Omega\left(\frac{1}{n}, T\right)]) + P_{\varphi^*}[\|\tilde{\varphi}'_T - \varphi^*\| > \frac{1}{n}] + P_{\varphi^*}[L_T^{(1)}(\tilde{\varphi}'_T) \neq 0]$

et $\limsup_{T \rightarrow +\infty} P_{\varphi^*}[\tilde{\varphi}'_T \neq \tilde{\varphi}_T] \leq 1 - \liminf_{T \rightarrow +\infty} P_{\varphi^*}[\Omega\left(\frac{1}{n}, T\right)]$.

Compte-tenu de (ii) lemme 2.2.3., il vient :

$$\lim_{T \rightarrow +\infty} P_{\varphi^*}[\tilde{\varphi}'_T \neq \tilde{\varphi}_T] = 0$$

d'où l'assertion (iv) du théorème.

b) Points (v) et (vi).

On a $\frac{1}{T} L_T^{(1)}(\tilde{\varphi}_T, \cdot) - \frac{1}{T} L_T^{(1)}(\varphi^*, \cdot) = \Gamma(\tilde{\varphi}_T - \varphi^*) + \tilde{\epsilon}_T(\tilde{\varphi}_T)(\tilde{\varphi}_T - \varphi^*)$ où,

d'après le lemme 2.2.2., pour tout $T > 0$ et $\eta > 0$:

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} \|\tilde{\epsilon}_T(\varphi)\| \leq \tilde{\epsilon}_{T, \eta}$$

avec $\lim_{\eta \rightarrow 0} \limsup_{T \rightarrow +\infty} P_{\varphi^*}[\tilde{\epsilon}_{T, \eta} \geq \rho] = 0$ pour tout $\rho > 0$.

Comme $[\|\tilde{\varphi}_T - \varphi^*\| < \eta] \cap [\tilde{\varepsilon}_T, \eta < \rho] \subset [\|\tilde{\varepsilon}_T(\tilde{\varphi}_T)\| < \rho]$

on a $P_{\varphi^*}[\|\tilde{\varepsilon}_T(\tilde{\varphi}_T)\| \geq \rho] \leq P_{\varphi^*}[\|\tilde{\varphi}_T - \varphi^*\| \geq \eta] + P_{\varphi^*}[\varepsilon_T, \eta \geq \rho]$

puis $\limsup_{T \rightarrow +\infty} P_{\varphi^*}[\|\tilde{\varepsilon}_T(\tilde{\varphi}_T)\| \geq \rho] \leq \limsup_{T \rightarrow +\infty} P_{\varphi^*}[\varepsilon_T, \eta \geq \rho]$

et enfin $\lim_{T \rightarrow +\infty} P_{\varphi^*}[\|\tilde{\varepsilon}_T(\tilde{\varphi}_T)\| \geq \rho] = 0$.

Ainsi

$$\|\Gamma^{-1} \left[\frac{1}{T^{1/2}} L_T^{(1)}(\tilde{\varphi}_T, \cdot) - \frac{1}{T^{1/2}} L_T^{(1)}(\varphi^*, \cdot) \right] - T^{1/2}(\tilde{\varphi}_T - \varphi^*)\| \leq \varepsilon_T \|T^{1/2}(\tilde{\varphi}_T - \varphi^*)\|$$

où $P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \varepsilon_T = 0$.

Alors, d'après P. BILLINGSLEY [12] théorème 10.1., on a :

$$P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \left\{ \Gamma^{-1} \left[\frac{1}{T^{1/2}} L_T^{(1)}(\tilde{\varphi}_T, \cdot) - \frac{1}{T^{1/2}} L_T^{(1)}(\varphi^*, \cdot) \right] - T^{1/2}(\tilde{\varphi}_T - \varphi^*) \right\} = 0.$$

En vertu de (ii) théorème 2.2.1. et de l'hypothèse (E3) on a donc :

$$\mathfrak{L}_{P_{\varphi^*}} \text{-} \lim_{T \rightarrow +\infty} T^{1/2}(\tilde{\varphi}_T - \varphi^*) = \Gamma^{-1} N(0, -\Gamma)$$

d'où l'assertion (v) du théorème.

Enfin, écrivant :

$$L_T(\varphi^*) - L_T(\tilde{\varphi}_T) = L_T^{(1)}(\tilde{\varphi}_T)(\varphi^* - \tilde{\varphi}_T) + \frac{1}{2} {}^t_{(\varphi^* - \tilde{\varphi}_T)} L_T^{(2)}(\tilde{\varphi}_T + \alpha(\tilde{\varphi}_T)(\varphi^* - \tilde{\varphi}_T))(\varphi^* - \tilde{\varphi}_T)$$

où $\alpha(\tilde{\varphi}_T) \in [0, 1]$, il vient :

$$P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \left\{ 2[L_T(\varphi^*) - L_T(\tilde{\varphi}_T)] - T^{1/2} {}^t_{(\varphi^* - \tilde{\varphi}_T)} \Gamma T^{1/2}(\varphi^* - \tilde{\varphi}_T) \right\} = 0$$

et par suite,

$$\mathfrak{L}_{P_{\varphi^*}} \text{-} \lim_{T \rightarrow +\infty} 2[L_T(\varphi^*) - L_T(\tilde{\varphi}_T)] = {}^t N(0, -\Gamma^{-1}) \Gamma N(0, -\Gamma^{-1})$$

d'où l'assertion (vi) du théorème. ■

2.2.5. Remarques.

a) La démonstration de la proposition 2.2.1. ne prend pas en compte le fait que la fonction $L_T(\varphi, \cdot)$ soit la fonction de Log-vraisemblance. Aussi l'énoncé de ce théorème reste-t-il vrai pour toute famille d'applications $\{\mathfrak{L}_T(\varphi, \cdot); T > 0, \varphi \in \Phi\}$ telle que pour tout $T > 0$, $\mathfrak{L}_T(\varphi, \cdot)$ soit \mathcal{A}_T -mesurable et $\mathfrak{L}_T(\cdot, \omega)$ soit deux fois continûment différentiable sur Φ . Il peut alors être intéressant de comparer la famille d'estimateurs $\{\tilde{\varphi}_T; T > 0\}$, basée sur $L_T(\varphi, \cdot)$, dite "famille d'estimateurs du maximum de vraisemblance", à celle basée sur $\mathfrak{L}_T(\varphi, \cdot)$ soit $\{\hat{\varphi}_T; T > 0\}$. Ainsi, par exemple, sous l'hypothèse qu'il existe une famille $\{\gamma_T; T > 0\}$ de nombres positifs telle que $\lim_{T \rightarrow +\infty} \gamma_T = 0$ et pour $i = 1, 2$ et tout $K > 0$ il existe une famille $\{\zeta_T^i(K); T > 0\}$ de variables aléatoires positives telle que :

$\{\zeta_T^i(K); T > 0\}$ est bornée en P_{φ^*} -probabilité i.e. $\lim_{A \rightarrow +\infty} \limsup_{T \rightarrow +\infty} P_{\varphi^*}[\zeta_T^i(K) > A] = 0$

et

$$\sup_{\|\varphi\| \leq K} \|\mathfrak{L}_T^{(i)}(\varphi, \cdot) - L_T^{(i)}(\varphi, \cdot)\| \leq \gamma_T \zeta_T^i(K)$$

il est facile de montrer que la famille $\{\gamma_T^{-1} \|\hat{\varphi}_T - \tilde{\varphi}_T\|; T > 0\}$ est bornée en P_{φ^*} -probabilité.

b) Si pour tout $T > 0$, $\omega \in \Omega$, l'équation :

$$L_T^{(1)}(\varphi, \omega) = 0$$

admet une solution unique $\bigvee \varphi_T(\omega)$ telle que :

$$L_T(\bigvee \varphi_T(\omega), \omega) = \text{Max}_{\varphi \in \Phi} L_T(\varphi, \omega)$$

la famille d'estimateurs fournie par la proposition 2.2.1. n'est évidemment rien d'autre que la famille $\{\bigvee \varphi_T; T > 0\}$ elle-même.

2.3. Cas de l'observation en temps discret.

On considère une structure statistique :

$$(\Omega, \mathcal{A}_T, \{P_{\varphi}^T; \varphi \in \Phi\})$$

liée à l'observation en temps continu sur $[0, T]$ d'une fonction aléatoire. On suppose que, $\mathfrak{L}_T(\dots)$ étant une application de $\Phi \times \Omega$ dans \mathbb{R} telle que, pour tout $\varphi \in \Phi$, $\mathfrak{L}_T(\varphi, \cdot)$ est \mathcal{A}_T -mesurable et, pour tout $\omega \in \Omega$, $\mathfrak{L}_T(\cdot, \omega)$ est deux fois continûment différentiable sur Φ , on peut construire un estimateur $\hat{\varphi}_T$ de φ par

$$\mathfrak{L}_T^{(1)}(\hat{\varphi}_T(\omega), \omega) = 0$$

avec de plus :

$\mathfrak{L}_T^{(2)}(\hat{\varphi}_T(\omega), \omega)$ est une matrice définie négative.

On se donne alors $\{\Pi_N\}_{N \geq 1}$ une suite de subdivisions de $[0, T]$, $\Pi_N = \{0 = t_0^N < t_1^N \dots < t_{K_N}^N = T\}$ telle que, posant $\delta_N = \max_{i=1, K_N} (t_i^N - t_{i-1}^N)$, on ait $\lim_{N \rightarrow +\infty} \delta_N = 0$. On désigne par \mathcal{A}_T^N la sous-tribu de \mathcal{A}_T liée à l'observation de la fonction aléatoire aux instants $t_0^N, t_1^N, \dots, t_{K_N}^N$. On considère une famille $\{\mathfrak{L}_{T,N}(\dots); T > 0, N \geq 1\}$ d'applications de $\Phi \times \Omega$ dans \mathbb{R} , telles que, pour tout $\varphi \in \Phi$, $\mathfrak{L}_{T,N}(\varphi, \cdot)$ est \mathcal{A}_T^N -mesurable et, pour tout $\omega \in \Omega$, $\mathfrak{L}_{T,N}(\cdot, \omega)$ est deux fois continûment différentiable sur Φ .

On énonce alors une proposition qui généralise un résultat présenté dans A. LE BRETON [33] :

2.3.1. PROPOSITION : "On suppose qu'il existe une suite $\{\gamma_N; N \geq 1\}$ de nombres positifs convergeant vers zéro et que pour tout $K > 0$ et $i = 1$ ou 2 il existe une suite $\{\nabla_N^i(K); N \geq 1\}$ de variables aléatoires positives \mathcal{A}_T -mesurables telles que :

$$(F 1) \quad \{\nabla_N^i(K); N \geq 1\} \text{ est bornée en } P_{\varphi^*}^T \text{-probabilité i.e.}$$

$$\lim_{A \rightarrow +\infty} \limsup_{N \rightarrow +\infty} P_{\varphi^*}[\nabla_N^i(K) > A] = 0.$$

$$(F 2) \quad \sup_{\|\varphi\| \leq K} \|\mathfrak{L}_{T,N}^{(i)}(\varphi, \cdot) - \mathfrak{L}_T^{(i)}(\varphi, \cdot)\| \leq \gamma_N \nabla_N^i(K).$$

Alors il existe une suite $\{\hat{\varphi}_{N,T}; N \geq 1\}$ d'estimateurs de φ vérifiant :

- (j) $\hat{\varphi}_{N,T}$ est \mathcal{A}_T -mesurable pour tout $N \geq 1$,
- (jj) $\lim_{N \rightarrow +\infty} P_{\varphi^*}^T[\mathcal{L}_{T,N}^{(1)}(\hat{\varphi}_{N,T}, \cdot) = 0] = 1$,
- (jjj) $P_{\varphi^*}^T\text{-}\lim_{N \rightarrow +\infty} \hat{\varphi}_{N,T} = \hat{\varphi}_T$.

En outre, si $\{\hat{\varphi}'_{N,T}; N \geq 1\}$ est une autre suite d'estimateurs vérifiant les conditions (j), (jj) et (jjj), alors :

- (jv) $\lim_{N \rightarrow +\infty} P_{\varphi^*}^T[\hat{\varphi}'_{N,T} = \hat{\varphi}_{N,T}] = 1$.

Enfin, si $\{\hat{\varphi}_{N,T}; N \geq 1\}$ est une famille d'estimateurs vérifiant (j), (jj) et (jjj) on a :

- (v) $\{\gamma_N^{-1} \|\hat{\varphi}_{N,T} - \hat{\varphi}_T\|; N \geq 1\}$ est une suite de variables aléatoires bornée en $P_{\varphi^*}^T$ -probabilité."

La démonstration de cette proposition suit des étapes analogues à celles de la démonstration de la proposition 2.2.1. Pour simplifier les notations nous supprimerons, dans le lemme suivant et dans les démonstrations l'indice T .

2.3.2. LEMME : "Il existe une famille $\{\Omega_N(K); K > 1\}$ d'événements vérifiant :

- (j) $\Omega_N(K) \in \mathcal{A}$
- (jj) $\lim_{K \rightarrow +\infty} \liminf_{N \rightarrow +\infty} P_{\varphi^*}[\Omega_N(K)] = 1$
telle que, pour tout $K > 0$, tout $N \geq 1$ et tout $\omega \in \Omega_N(K)$ l'équation $\mathcal{L}_N^{(1)}(\varphi, \omega) = 0$ admette une solution unique $\hat{\varphi}_N(\omega)$ dans la boule $\|\varphi - \hat{\varphi}_N(\omega)\| \leq \frac{1}{K}$, cette solution vérifiant de plus $\mathcal{L}_N(\hat{\varphi}_N(\omega), \omega) > \mathcal{L}_N(\varphi, \omega)$ pour tout $\varphi \neq \hat{\varphi}_N(\omega)$ tel que $\|\varphi - \hat{\varphi}_N(\omega)\| \leq 1/K$."

Démonstration : On écrit, pour $\omega \in \Omega$, $u \in \mathbb{R}^d$:

$$\mathcal{L}_N^{(1)}(\hat{\varphi}_N(\omega) + u, \omega) = \mathcal{L}_N^{(1)}(\hat{\varphi}_N(\omega), \omega) + \mathcal{L}_N^{(2)}(\hat{\varphi}_N(\omega), \omega)u + \{\mathcal{L}_N^{(2)}(\varphi'(\omega, u), \omega) - \mathcal{L}_N^{(2)}(\hat{\varphi}_N(\omega), \omega)\}u$$

où $\varphi'(w, u) = \hat{\varphi}(w) + \lambda(\hat{\varphi}(w), u)u$; $\lambda(\hat{\varphi}(w), u) \in [0, 1]$.

Soit $K > 0$ donné. L'application $\mathfrak{L}_N^{(1)}(\hat{\varphi}(w) + \cdot, w)$ est continue sur $\|u\| \leq \frac{1}{K}$. De plus :

$$\langle u, \mathfrak{L}_N^{(1)}(\hat{\varphi}(w) + u, w) \rangle = \langle u, \mathfrak{L}_N^{(1)}(\hat{\varphi}(w), w) + {}^t u \mathfrak{L}_N^{(2)}(\hat{\varphi}(w), w)u + {}^t u \{ \mathfrak{L}_N^{(2)}(\varphi'(w, u), w) - \mathfrak{L}_N^{(2)}(\hat{\varphi}(w), w) \} \rangle$$

Compte-tenu du fait que $[-\mathfrak{L}_N^{(2)}(\hat{\varphi}(w), w)]$ est une matrice définie positive, on a :

$$\inf_{\|u\|=1} \{ -{}^t u \mathfrak{L}_N^{(2)}(\hat{\varphi}(w), w)u \} = q(w) > 0$$

et
$$-{}^t u \mathfrak{L}_N^{(2)}(\hat{\varphi}(w), w)u \geq q(w) \|u\|^2 .$$

Ainsi $\langle u, \mathfrak{L}_N^{(1)}(\hat{\varphi}(w) + u, w) \rangle$ sera strictement négatif pour $\|u\| = \frac{1}{K}$ si

$$\|\mathfrak{L}_N^{(1)}(\hat{\varphi}(w), w)\| < \frac{q(w)}{2K} \quad \text{et} \quad \sup_{\|u\| \leq \frac{1}{K}} \|\mathfrak{L}_N^{(2)}(\varphi'(w, u), w) - \mathfrak{L}_N^{(2)}(\hat{\varphi}(w), w)\| < \frac{q(w)}{2} .$$

On considère alors, pour $K > 1$, l'événement \mathcal{A} -mesurable $\Omega_N(K)$ défini par :

$$\Omega_N(K) = \{ \omega \in \Omega : \gamma_{N \nabla N}^1(K - \frac{1}{K})(\omega) < \frac{q(\omega)}{2K} , \gamma_{N \nabla N}^2(K)(\omega) < \frac{q(\omega)}{4} , \|\hat{\varphi}(\omega)\| \leq K - \frac{1}{K} \} .$$

Pour tout $\omega \in \Omega_N(K)$ on a :

$$\sup_{\|\varphi\| \leq K - \frac{1}{K}} \|\mathfrak{L}_N^{(1)}(\varphi, \omega) - \mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega), \omega)\| < \frac{q(\omega)}{2K}$$

et, compte-tenu de $\|\hat{\varphi}(\omega)\| \leq K - \frac{1}{K}$, on a aussi :

$$\|\mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega), \omega) - \mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega), \omega)\| < \frac{q(\omega)}{2K} .$$

De plus
$$\sup_{\|\varphi\| \leq K} \|\mathfrak{L}_N^{(2)}(\varphi, \omega) - \mathfrak{L}_N^{(2)}(\hat{\varphi}(\omega), \omega)\| < \frac{q(\omega)}{4}$$

et, compte-tenu de ce que $\|\varphi'(w, u)\| \leq K$ pour $\|u\| \leq \frac{1}{K}$ et de ce que $\|\hat{\varphi}(w)\| \leq K$, on a :

$$\sup_{\|u\| \leq \frac{1}{K}} \|\mathfrak{L}_N^{(2)}(\varphi'(w, u), w) - \mathfrak{L}_N^{(2)}(\hat{\varphi}(w), w)\| < \frac{q(w)}{2} .$$

En définitive, pour $\omega \in \Omega_N(K)$ l'application $\mathfrak{L}_N^{(1)}(\hat{\varphi}(w) + \cdot, w)$ est continue sur $\|u\| \leq \frac{1}{K}$ et pour $\|u\| = \frac{1}{K}$,

$$\langle u, \mathfrak{L}_N^{(1)}(\hat{\varphi}(w) + u, w) \rangle < 0 .$$

Le lemme 2.1.1. assure donc que l'équation :

$$\mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega) + u, \omega) = 0$$

admet une solution $X_N(\omega)$ vérifiant $\|X_N(\omega)\| < \frac{1}{K}$.

Considérons :

$$\|\mathfrak{L}_N^{(2)}(\hat{\varphi}(\omega) + u, \omega) - \mathfrak{L}_N^{(2)}(\hat{\varphi}(\omega), \omega)\| < \frac{q(\omega)}{2}$$

pour $\omega \in \Omega_N(K)$ et $\|u\| \leq \frac{1}{K}$, on voit que $\mathfrak{L}_N^{(2)}(\hat{\varphi}(\omega) + u, \omega)$ est définie négative. Calculant, pour $\|u\| \leq \frac{1}{K}$,

$$\mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega) + u, \omega) - \mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega) + X_N(\omega), \omega) = \int_0^1 \frac{d}{d\alpha} [\mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega) + X_N(\omega) + \alpha(u - X_N(\omega)), \omega)] d\alpha$$

c'est-à-dire :

$$\mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega) + u, \omega) = \int_0^1 \mathfrak{L}_N^{(2)}(\hat{\varphi}(\omega) + X_N(\omega) + \alpha(u - X_N(\omega)), \omega) (u - X_N(\omega)) d\alpha$$

il vient :

$$\begin{aligned} \langle \mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega) + u, \omega), u - X_N(\omega) \rangle &= \int_0^1 \langle (u - X_N(\omega)) \mathfrak{L}_N^{(2)}(\hat{\varphi}(\omega) + X_N(\omega) + \alpha(u - X_N(\omega)), \omega) (u - X_N(\omega)) \rangle d\alpha \\ &< 0 \quad \text{si } u \neq X_N(\omega). \end{aligned}$$

Ainsi, pour $\|u\| \leq \frac{1}{K}$ et $u \neq X_N(\omega)$ on a $\mathfrak{L}_N^{(1)}(\hat{\varphi}(\omega) + u, \omega) \neq 0$.

En outre, $\mathfrak{L}_N(\hat{\varphi}(\omega) + \cdot, \omega)$ atteint évidemment son maximum sur $\|u\| \leq \frac{1}{K}$ en $X_N(\omega)$.

Le résultat du lemme s'acquiert en posant $\hat{\varphi}_N(\omega) = \hat{\varphi}(\omega) + X_N(\omega)$ sous réserve de vérification de (jj).

Or, considérons l'inégalité :

$$P_{\varphi^*}[\bigcup_N \Omega_N(K)] \leq P_{\varphi^*}[\gamma_N \nabla_N^1(K - \frac{1}{K}) \geq \frac{q(\cdot)}{2K}] + P_{\varphi^*}[\gamma_N \nabla_N^2(K) \geq \frac{q(\cdot)}{4}] + P_{\varphi^*}[\|\hat{\varphi}\| > K - \frac{1}{K}].$$

En vertu de l'hypothèse (F1) on a :

$$\lim_{N \rightarrow +\infty} P_{\varphi^*}[\gamma_N \nabla_N^1(K - \frac{1}{K}) \geq \frac{q(\cdot)}{2K}] = \lim_{N \rightarrow +\infty} P_{\varphi^*}[\gamma_N \nabla_N^2(K) \geq \frac{q(\cdot)}{2}] = 0.$$

Ainsi on a :

$$\limsup_{N \rightarrow +\infty} P_{\varphi^*} [\bigcap \Omega_N(K)] \leq P_{\varphi^*} [\| \hat{\varphi} \| > K - \frac{1}{K}]$$

et alors :

$$\lim_{K \rightarrow +\infty} \limsup_{N \rightarrow +\infty} P_{\varphi^*} [\bigcap \Omega_N(K)] = 0$$

ce qui achève la démonstration du lemme. ■

2.3.3. Démonstration de la proposition 2.3.1.

a) Points (j) jusqu'à (jv).

Posons $\Omega_N = \bigcup_{K \in \mathbb{N}^*} \Omega_N(K) \in \mathcal{A}$ et considérons $\omega \in \Omega_N$. D'après le lemme 2.3.2. il existe un entier $K(\omega) \geq 1$ tel que l'équation $\mathfrak{L}_N^{(1)}(\varphi, \omega) = 0$ admette une solution unique $\tilde{\varphi}_N(\omega)$ dans la boule $\| \varphi - \hat{\varphi}(\omega) \| \leq \frac{1}{K(\omega)}$.

Nous sommes donc dans les conditions d'application du lemme 2.1.2. en y faisant :

$$\begin{aligned} (\Omega, \mathcal{A}) &= (\Omega_N, \mathcal{A} \cap \Omega_N) \quad (\text{rappelons que, par convention d'écriture, dans le second membre } \mathcal{A} \text{ n'est autre que } \mathcal{A}_T). \\ \mathfrak{L}(\omega, u) &= \mathfrak{L}_N^{(1)}(u, \omega) \end{aligned}$$

$$\text{et } h(\omega) = \hat{\varphi}(\omega)$$

et par conséquent, l'application $\omega \rightsquigarrow \tilde{\varphi}_N(\omega)$ définie sur $(\Omega_N, \mathcal{A} \cap \Omega_N)$ est $\mathcal{A} \cap \Omega_N$ -mesurable.

$$\text{On pose alors } \hat{\varphi}_N(\omega) = \begin{cases} \tilde{\varphi}_N(\omega) & \text{si } \omega \in \Omega_N \\ \varphi_0 & \text{si } \omega \in \bigcap \Omega_N \end{cases}$$

où φ_0 est une valeur quelconque du paramètre.

L'application $\hat{\varphi}_N$ ainsi définie est \mathcal{A} -mesurable.

Soit $\eta > 0$ donné: pour tout $K \geq 1$ tel que $\frac{1}{K} < \eta$ on a :

$$\Omega_N(K) \subset [\|\hat{\varphi}_N - \hat{\varphi}\| < \eta] .$$

Comme pour tout $\varepsilon > 0$ il existe K_0 tel que $K \geq K_0$ implique que :

$$\liminf_{N \rightarrow +\infty} P_{\varphi^*}[\Omega_N(K)] > 1 - \varepsilon ,$$

choisissant K tel que $\frac{1}{K} < \eta$ et $K \geq K_0$, il existe N_0 tel que :

$$N > N_0 \Rightarrow P_{\varphi^*}[\Omega_N(K)] > 1 - \varepsilon$$

et alors :

$$P_{\varphi^*}[\|\hat{\varphi}_N - \hat{\varphi}\| < \eta] > 1 - \varepsilon .$$

Ainsi $P_{\varphi^*} - \lim_{N \rightarrow +\infty} \hat{\varphi}_N = \hat{\varphi}$ et le point (jjj) est démontré.

Maintenant, puisque $\Omega_N(K) \subset [\mathfrak{L}_N^{(1)}(\hat{\varphi}_N) = 0]$, si $N > N_0$, on a :

$$P_{\varphi^*}[\mathfrak{L}_N^{(1)}(\hat{\varphi}_N) = 0] \geq P_{\varphi^*}[\Omega_N(K)] > 1 - \varepsilon$$

et donc :

$$\lim_{N \rightarrow +\infty} P_{\varphi^*}[\mathfrak{L}_N^{(1)}(\hat{\varphi}_N) = 0] = 1$$

i.e. (jj) est acquis.

Enfin, considérons une autre famille de variables aléatoires $\{\hat{\varphi}'_N ; N \geq 1\}$ vérifiant les conditions (j) jusqu'à (jjj).

On a, pour tout $K \geq 1$,

$$\{\omega \in \Omega : \omega \in \Omega_N(K), \|\hat{\varphi}'_N(\omega) - \hat{\varphi}(\omega)\| \leq \frac{1}{K}, \mathfrak{L}_N^{(1)}(\hat{\varphi}'_N(\omega), \omega) = 0\} \subset [\hat{\varphi}'_N = \hat{\varphi}_N]$$

donc,

$$P_{\varphi^*}[\hat{\varphi}'_N \neq \hat{\varphi}_N] \leq (1 - P_{\varphi^*}[\Omega_N(K)]) + P_{\varphi^*}[\|\hat{\varphi}'_N - \hat{\varphi}\| > \frac{1}{K}] + P_{\varphi^*}[\mathfrak{L}_N^{(1)}(\hat{\varphi}'_N) \neq 0] .$$

Alors :

$$\limsup_{N \rightarrow +\infty} P_{\varphi^*}[\hat{\varphi}'_N \neq \hat{\varphi}_N] \leq 1 - \liminf_{N \rightarrow +\infty} P_{\varphi^*}[\Omega_N(K)] .$$

Compte-tenu de (jj) lemme 2.3.2., il vient :

$$\lim_{N \rightarrow +\infty} P_{\varphi^*}[\hat{\varphi}'_N \neq \hat{\varphi}_N] = 0$$

d'où l'assertion (jv) du théorème.

b) Point (v).

Soit $\{\hat{\varphi}_N; N \geq 1\}$ une suite de variables aléatoires vérifiant les conditions (j) jusqu'à (jv). Pour $\omega \in [\mathcal{L}_N^{(1)}(\hat{\varphi}_N) = 0]$ on peut écrire :

$$\begin{aligned} & \mathcal{L}_N^{(1)}(\hat{\varphi}_N(\omega), \omega) - \mathcal{L}_N^{(1)}(\hat{\varphi}(\omega), \omega) = \mathcal{L}_N^{(1)}(\hat{\varphi}_N(\omega), \omega) \\ & = \mathcal{L}_N^{(1)}(\hat{\varphi}(\omega), \omega) + \mathcal{L}_N^{(2)}(\hat{\varphi}(\omega) + \alpha_N(\omega)(\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega)))(\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega)) \\ & = \mathcal{L}_N^{(2)}(\hat{\varphi}(\omega) + \alpha_N(\omega)(\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega)))(\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega)) \end{aligned}$$

où $\alpha_N(\omega) \in [0, 1]$.

Alors pour $\omega \in [\Omega_N(K)] \cap [\|\hat{\varphi}_N - \hat{\varphi}\| \leq \frac{1}{K}] \cap [\|\hat{\varphi}_N\| \leq K - \frac{1}{K}] \cap [\mathcal{L}_N^{(1)}(\hat{\varphi}_N) = 0]$,

$$\begin{aligned} \gamma_N^{-1}(\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega)) &= \gamma_N^{-1}[\mathcal{L}_N^{(2)}(\hat{\varphi}(\omega) + \alpha_N(\omega)(\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega))]^{-1} \{ \mathcal{L}_N^{(1)}(\hat{\varphi}_N(\omega), \omega) - \mathcal{L}_N^{(1)}(\hat{\varphi}(\omega), \omega) \} \\ \text{et} \quad \gamma_N^{-1} \|\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega)\| &\leq \|[\mathcal{L}_N^{(2)}(\hat{\varphi}(\omega) + \alpha_N(\omega)(\hat{\varphi}_N(\omega) - \hat{\varphi}(\omega))]^{-1}\| \cdot \nabla_N^1(K). \end{aligned}$$

Par suite, pour $A > 0$ donné,

$$\begin{aligned} P_{\varphi^*}[\gamma_N^{-1} \|\hat{\varphi}_N - \hat{\varphi}\| > A] &\leq P_{\varphi^*}[\bigcup \Omega_N(K)] + P_{\varphi^*}[\|\hat{\varphi}_N - \hat{\varphi}\| > \frac{1}{K}] \\ &\quad + P_{\varphi^*}[\|\hat{\varphi}_N\| > K - \frac{1}{K}] + P_{\varphi^*}[\mathcal{L}_N^{(1)}(\hat{\varphi}_N) \neq 0] \\ &\quad + P_{\varphi^*}[\|[\mathcal{L}_N^{(2)}(\hat{\varphi} + \alpha_N(\cdot)(\hat{\varphi}_N - \hat{\varphi}))^{-1}\| \cdot \nabla_N^1(K) > A]. \end{aligned}$$

Soit alors $\varepsilon > 0$ donné : il existe $K_0 > 0$ tel que :

$$P_{\varphi^*}[\|\hat{\varphi}\| > K_0 - \frac{1}{K_0}] < \frac{\varepsilon}{10}$$

(et par suite, d'après la démonstration du lemme 2.3.2.,

$$\limsup_{N \rightarrow +\infty} P_{\varphi^*}[\bigcup \Omega_N(K_0)] < \frac{\varepsilon}{5},$$

$$\text{et} \quad \lim_{N \rightarrow +\infty} P_{\varphi^*}[\|\hat{\varphi}_N\| > K_0 - \frac{1}{K_0}] = P_{\varphi^*}[\|\hat{\varphi}\| > K_0 - \frac{1}{K_0}].$$

On peut donc choisir N_1 tel que :

$$N \geq N_1 \Rightarrow |P_{\varphi^*}[\|\hat{\varphi}_N\| > K_0 - \frac{1}{K_0}] - P_{\varphi^*}[\|\hat{\varphi}\| > K_0 - \frac{1}{K_0}]| < \frac{\varepsilon}{10}$$

de sorte que
$$P_{\varphi^*}[\|\hat{\varphi}_N\| > K_0 - \frac{1}{K_0}] < \frac{\varepsilon}{5} .$$

En outre, utilisant le fait que :

$$\lim_{N \rightarrow +\infty} P_{\varphi^*}[\mathcal{L}_N^{(1)}(\hat{\varphi}_N) \neq 0] = 0 \quad (\text{resp. } \{ \|\mathcal{L}^{(2)}(\hat{\varphi} + \alpha_N(\cdot)(\hat{\varphi}_N - \hat{\varphi}))^{-1}\|_{\nabla_N^1(K_0)}; N \geq 1 \})$$

est bornée en P_{φ^*} -probabilité),

on peut trouver N_2 (resp. A et N_3) tel (resp. tels) que :

$$N \geq N_2 \Rightarrow P_{\varphi^*}[\mathcal{L}_N^{(1)}(\hat{\varphi}_N) \neq 0] < \frac{\varepsilon}{5}$$

$$(\text{resp. } N \geq N_3 \Rightarrow P_{\varphi^*}[\|\mathcal{L}^{(2)}(\hat{\varphi} + \alpha(\cdot)(\hat{\varphi}_N - \hat{\varphi}))^{-1}\|_{\nabla_N^1(K_0)} > A] < \frac{\varepsilon}{5}) .$$

Enfin, comme :

$$\limsup_{N \rightarrow +\infty} P_{\varphi^*}[\bigcup_N \Omega_N(K_0)] < \frac{\varepsilon}{5}$$

et que
$$\lim_{N \rightarrow +\infty} P_{\varphi^*}[\|\hat{\varphi}_N - \hat{\varphi}\| > \frac{1}{K_0}] = 0 ,$$

il existe N_4 tel que $N \geq N_4 \Rightarrow P_{\varphi^*}[\bigcup_N \Omega_N(K_0)] < \frac{\varepsilon}{5}$ et $P_{\varphi^*}[\|\hat{\varphi}_N - \hat{\varphi}\| > \frac{1}{K_0}] < \frac{\varepsilon}{5}$.

En définitive, $\varepsilon > 0$ étant donné, il existe $N_0 = \text{Max}(N_i; i=1,4)$ et $A > 0$ tels que pour $N \geq N_0$ on ait :

$$P_{\varphi^*}[\gamma_N^{-1} \|\hat{\varphi}_N - \hat{\varphi}\| > A] < \varepsilon .$$

Alors
$$\lim_{A \rightarrow +\infty} \limsup_{N \rightarrow +\infty} P_{\varphi^*}[\gamma_N^{-1} \|\hat{\varphi}_N - \hat{\varphi}\| > A] = 0$$

ce qui achève la démonstration. ■

2.3.4. Remarques.

a) La proposition 2.3.1. n'assure par l'existence d'une suite $\{\tilde{\varphi}_{N,T}; N \geq 1\}$ d'estimateurs vérifiant les conditions (j), (jj), (jjj) et (v) telle que de plus, pour tout $N \geq 1$, $\tilde{\varphi}_{N,T}$ soit \mathcal{O}_T^N -mesurable. Toutefois si une telle suite existe on sait que :

$$\lim_{N \rightarrow +\infty} P_{\varphi^*}^T [\tilde{\varphi}_{N,T} = \hat{\varphi}_{N,T}] = 1 .$$

En particulier si, en plus des hypothèses (F1) et (F2), pour tout $N \geq 1$ et $\omega \in \Omega$, l'équation $\mathfrak{L}_{T,N}(\varphi, \omega) = 0$ admet une solution unique $\tilde{\varphi}_{N,T}$, l'estimateur $\tilde{\varphi}_{N,T}$ est \mathcal{O}_T^N -mesurable et vérifie les conditions (jjj) et (v).

b) Supposons que la famille d'estimateurs $\{\hat{\varphi}_{N,T}; N \geq 1\}$ fournie par la proposition 2.3.1. est adaptée à $(\mathcal{O}_T^N; N \geq 1)$. Considérons $\{\Lambda_{T,N}(\dots); N \geq 1\}$ une autre famille d'applications vérifiant les mêmes hypothèses que $\mathfrak{L}_T(\dots)$ en remplaçant, pour tout $N \geq 1$, \mathcal{O}_T par \mathcal{O}_T^N .

On suppose que pour tout $K > 0$ et $i = \overline{1,2}$ il existe une suite $\{\zeta_N^i(K); N \geq 1\}$ de variables aléatoires positives telles que :

(F1)' pour tout $N \geq 1$, $\zeta_N^i(K)$ est \mathcal{O}_T^N -mesurable ;

(F2)' $\{\zeta_N^i(K); N \geq 1\}$ est bornée en $P_{\varphi^*}^T$ -probabilité ;

(F3)' $\sup_{\|\varphi\| \leq K} \|\mathfrak{L}_{T,N}^{(i)}(\varphi, \cdot) - \Lambda_{T,N}^{(i)}(\varphi, \cdot)\| \leq \gamma_N \zeta_N^i(K)$.

Il est alors clair que la famille $\{\Lambda_{T,N}(\dots); N \geq 1\}$ vérifie aussi les hypothèses (F1) et (F2) et, par suite, la proposition 2.3.1. fournit une suite $\{\check{\varphi}_{N,T}; N \geq 1\}$ d'estimateurs vérifiant les conclusions de la dite proposition. Il est alors facile de montrer que $\{\check{\varphi}_{N,T}; N \geq 1\}$ peut être choisie de manière à ce que, pour tout $N \geq 1$, $\check{\varphi}_{N,T}$ soit \mathcal{O}_T^N -mesurable et que de plus $\{\gamma_N^{-1} \|\hat{\varphi}_{N,T} - \check{\varphi}_{N,T}\|; N \geq 1\}$ soit bornée en $P_{\varphi^*}^T$ -probabilité.

c) L'énoncé de la proposition 2.3.1. ne prend pas en compte le fait que les fonctions \mathfrak{L}_T et $\mathfrak{L}_{T,N}$ soient ou non les fonctions de Log-vraisemblance pour observation en temps continu et en temps discret respectivement. Si c'est le cas on dira que les estimateurs sont "de maximum de vraisemblance".

En vu d'appliquer les propositions précédentes dans le cadre du problème posé, on étudie, dans le paragraphe suivant, quelques propriétés des intégrales stochastiques.

§ 3. - QUELQUES PROPRIETES DES INTEGRALES STOCHASTIQUES. DERIVABILITE PAR RAPPORT A UN PARAMETRE. NORMALITE ASYMPTOTIQUE.

3.1. Dérivabilité par rapport à un paramètre.

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ une base de processus, $W = (W_t ; t \geq 0)$ un processus de mouvement brownien p-dimensionnel $\{\psi_\varphi = (\psi_\varphi(t) ; t \geq 0) , \varphi \in \Phi\}$ (resp. $\{\Lambda_\varphi(t) ; t \geq 0\} , \varphi \in \Phi\}$) une famille de fonctions aléatoires mesurables à valeurs dans $\mathcal{L}(\mathbb{R}^p ; \mathbb{R}^q)$ (resp. \mathbb{R}^q) définis sur cette base, tels que, pour $T > 0$ donné, on ait, pour tout $\varphi \in \Phi$:

$$P\left[\int_0^T (\|\psi_\varphi(t)\|^2 dt + \|\Lambda_\varphi(t)\|^2 dt) < +\infty\right] = 1 .$$

Alors, pour tout $\varphi \in \Phi$, l'intégrale :

$$I_\varphi(T) = \int_0^T \psi_\varphi(t) dW_t + \int_0^T \Lambda_\varphi(t) dt$$

est bien définie.

On trouve dans [23] I.I. GIHMAN et A.V. SKOROHOD des résultats concernant la dérivabilité en moyenne quadratique de la solution d'une équation différentielle stochastique dont les coefficients dépendent d'un paramètre. Ici nous nous intéressons, en vue de définir des versions régulières de fonctions de vraisemblance, à l'existence de versions presque sûrement continûment dérivables par rapport à φ du processus $(\varphi, \omega) \rightsquigarrow I_\varphi(T)(\omega)$.

On a le résultat suivant :

3.1.1. PROPOSITION : "On suppose que, pour tout $(t, \omega) \in [0, T] \times \Omega$, les applications $\varphi \rightsquigarrow \psi_\varphi(t, \omega)$ et $\varphi \rightsquigarrow \Lambda_\varphi(t, \omega)$ sont deux fois différentiables sur Φ et que les dérivées partielles premières $\frac{\partial}{\partial \varphi_i} \psi_\varphi(t, \omega)$, $\frac{\partial}{\partial \varphi_i} \Lambda_\varphi(t, \omega)$, $i = \overline{1, d}$ et secondes $\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} \psi_\varphi(t, \omega)$, $\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} \Lambda_\varphi(t, \omega)$, $i, j = \overline{1, d}$ vérifient respectivement les conditions :

$$(G1) \quad \lim_{\Delta \varphi_i \rightarrow 0} E \left\{ \int_0^T \left\| \frac{\psi_{\varphi + \Delta \varphi_i}(t) - \psi_\varphi(t)}{\Delta \varphi_i} - \frac{\partial}{\partial \varphi_i} \psi_\varphi(t) \right\|^2 dt \right\} = 0$$

$$\lim_{\Delta\varphi_i \rightarrow 0} E \left\{ \int_0^T \left\| \frac{\Lambda_{\varphi+\Delta\varphi_i}(t) - \Lambda_{\varphi}(t)}{\Delta\varphi_i} - \frac{\partial}{\partial\varphi_i} \psi_{\varphi}(t) \right\|^2 dt \right\} = 0$$

où si ${}^t\varphi = (\varphi_1, \dots, \varphi_d)$ et $\Delta\varphi_i \in \mathbb{R}$ on pose
 ${}^t(\varphi + \Delta\varphi_i) = (\varphi_1, \dots, \varphi_{i-1}, \varphi_i + \Delta\varphi_i, \varphi_{i+1}, \dots, \varphi_d)$,

(G2) il existe un processus $(U_t; t \in [0, T])$ tel que
 $E \int_0^T |U_t|^r dt \leq K < +\infty$ pour un réel $r \geq 2$, $r > d$ et :

$$\sup_{\varphi \in \Phi} \left\| \frac{\partial}{\partial\varphi_i \partial\varphi_j} \psi_{\varphi}(t, \omega) \right\| \leq U_t(\omega)$$

$$(t, \omega) \in [0, T] \times \Omega.$$

$$\sup_{\varphi \in \Phi} \left\| \frac{\partial}{\partial\varphi_i \partial\varphi_j} \Lambda_{\varphi}(t, \omega) \right\| \leq U_t(\omega)$$

On définit alors les processus $\varphi \rightsquigarrow I_{\varphi}(T) = \int_0^T \psi_{\varphi}(t) dW_t + \int_0^T \Lambda_{\varphi}(t) dt$ et
 $\varphi \rightsquigarrow I_{\varphi}^{(1)}(T) = \left(\int_0^T \frac{\partial}{\partial\varphi_i} \psi_{\varphi}(t) dW_t + \int_0^T \frac{\partial}{\partial\varphi_i} \Lambda_{\varphi}(t) dt ; i = \overline{1, d} \right) = (I_{\varphi, i}^{(1)}(T) ; i = \overline{1, d})$.

Il existe des fonctions aléatoires $\varphi \rightsquigarrow \hat{I}_{\varphi}(T)$ et $\varphi \rightsquigarrow \hat{I}_{\varphi}^{(1)}(T)$ à trajectoires
continues telles que :

- (i) $\hat{I}_{\varphi}(T) = I_{\varphi}(T)$ et $\hat{I}_{\varphi}^{(1)}(T) = I_{\varphi}^{(1)}(T)$ P-presque sûrement
pour tout $\varphi \in \Phi$;
- (ii) $\hat{I}_{\varphi, i}^{(1)}(T)(\omega) = \frac{\partial}{\partial\varphi_i} \hat{I}_{\varphi}(T)(\omega)$ pour $\omega \in \Omega$, $i = \overline{1, d}$."

La démonstration de cette proposition s'appuie sur les préliminaires suivants :

3.1.2. LEMME (voir par exemple [40] J. NEVEU, p. 6-6) :

"Soit $(M_t)_{t \geq 0}$ une martingale réelle continue, nulle à l'origine, de processus croissant $(A_t; t \geq 0)$. Alors si $r \geq 2$, on a :

$$E \left[\sup_{s \leq t} |M_s| \right] \leq c_r E[A_t^{r/2}]$$

où c_r est une constante ne dépendant que de r ."

3.1.3. LEMME (voir par exemple [40] J. NEVEU, p. 6-5) :
 "Soit $(Z_x, x \in \mathbb{R}^d)$ une collection de variables aléatoires définies sur (Ω, \mathcal{F}, P) telle que :

$$E[\|Z_x - Z_y\|^r] \leq a \|x-y\|^{d+\varepsilon} ; \quad x, y \in \mathbb{R}^d$$

pour une constante finie a et deux réels $r \geq 1$, $\varepsilon > 0$. Alors il existe une fonction aléatoire Z^* définie sur $\Omega \times \mathbb{R}^d$ à trajectoires continues telle que :

$$Z_x^* = Z_x \quad P\text{-presque sûrement pour tout } x \in \mathbb{R}^d . "$$

3.1.4. Démonstration de la proposition 3.1.1.

On montre d'abord que $\varphi \rightsquigarrow I_\varphi(T)$ est différentiable en moyenne quadratique de dérivée $I_\varphi^{(1)}(T)$. On évalue :

$$\begin{aligned} & E \left\| \frac{I_{\varphi+\Delta\varphi_i}(T) - I_\varphi(T)}{\Delta\varphi_i} - I_{\varphi,i}^{(1)}(T) \right\|^2 \\ &= E \left\| \int_0^T \left[\frac{\psi_{\varphi+\Delta\varphi_i}(t) - \psi_\varphi(t)}{\Delta\varphi_i} - \frac{\partial}{\partial\varphi_i} \psi_\varphi(t) \right] dW_t + \int_0^T \left[\frac{\Lambda_{\varphi+\Delta\varphi_i}(t) - \Lambda_\varphi(t)}{\Delta\varphi_i} - \frac{\partial}{\partial\varphi_i} \Lambda_\varphi(t) \right] dt \right\|^2 \\ &\leq 2 \left\{ \int_0^T E \left\| \frac{\psi_{\varphi+\Delta\varphi_i}(t) - \psi_\varphi(t)}{\Delta\varphi_i} - \frac{\partial}{\partial\varphi_i} \psi_\varphi(t) \right\|^2 dt \right. \\ &\quad \left. + \int_0^T E \left\| \frac{\Lambda_{\varphi+\Delta\varphi_i}(t) - \Lambda_\varphi(t)}{\Delta\varphi_i} - \frac{\partial}{\partial\varphi_i} \Lambda_\varphi(t) \right\|^2 dt \right\} . \end{aligned}$$

Alors d'après l'hypothèse (G1) on a :

$$\lim_{\Delta\varphi_i \rightarrow 0} E \left\| \frac{I_{\varphi+\Delta\varphi_i}(T) - I_\varphi(T)}{\Delta\varphi_i} - I_{\varphi,i}^{(1)}(T) \right\|^2 = 0 .$$

D'autre part, r étant supérieur à 2 :

$$\begin{aligned}
& E \left\| \int_0^T \frac{\partial}{\partial \varphi_i} \psi_\varphi(t) dW_t - \int_0^T \frac{\partial}{\partial \varphi_i} \psi_{\varphi'}(t) dW_t \right\|^r \\
& \leq C_r E \left[\left(\int_0^T \left\| \frac{\partial}{\partial \varphi_i} \psi_\varphi(t) - \frac{\partial}{\partial \varphi_i} \psi_{\varphi'}(t) \right\|^2 dt \right)^{r/2} \right] \quad (\text{lemme 3.1.2}) \\
& \leq C_r T^{\frac{r-2}{2}} E \left[\int_0^T \left\| \frac{\partial}{\partial \varphi_i} \psi_\varphi(t) - \frac{\partial}{\partial \varphi_i} \psi_{\varphi'}(t) \right\|^r dt \right] \quad (\text{Hölder}) \\
& \leq C_r T^{\frac{r-2}{2}} K \|\varphi - \varphi'\|^r \quad (\text{hypothèse G2})
\end{aligned}$$

et

$$\begin{aligned}
E \left\| \int_0^T \frac{\partial}{\partial \varphi_i} \Lambda_\varphi(t) dt - \int_0^T \frac{\partial}{\partial \varphi_i} \Lambda_{\varphi'}(t) dt \right\|^r & \leq T^{r-1} E \left(\int_0^T \left\| \frac{\partial}{\partial \varphi_i} \Lambda_\varphi(t) - \frac{\partial}{\partial \varphi_i} \Lambda_{\varphi'}(t) \right\|^r dt \right) \quad (\text{Hölder}) \\
& \leq T^{r-1} K \|\varphi - \varphi'\|^r \quad (\text{hypothèse G2}).
\end{aligned}$$

Compte-tenu de l'inégalité $|a+b|^r \leq 2^{r-1} [|a|^r + |b|^r]$, il vient :

$$E \left\| I_{\varphi,i}^{(1)}(T) - I_{\varphi',i}^{(1)}(T) \right\|^r \leq 2^{r-1} \left\{ C_r T^{\frac{r-2}{2}} + T^{r-1} K \right\} \|\varphi - \varphi'\|^r.$$

En définitive $\varphi \rightsquigarrow I_\varphi(T)$ admet, au sens de la moyenne quadratique, des dérivées partielles continues : cette application est donc continûment différentiable en moyenne quadratique sur Φ , de différentielle définie par

$$dI_\varphi(T)(h) = \sum_{i=1}^d I_{\varphi,i}^{(1)}(T) \cdot h_i \quad \text{où } h = {}^t(h_1, \dots, h_d) \in \mathbb{R}^d.$$

En outre, puisque $r > d$, d'après le lemme 3.1.3, il existe une fonction aléatoire $(\varphi, \omega) \rightsquigarrow \tilde{I}_\varphi^{(1)}(T)(\omega) = (\tilde{I}_{\varphi,i}^{(1)}(T)(\omega); i=\overline{1,d})$ à trajectoires continues telle que :

$$\tilde{I}_\varphi^{(1)}(T) = I_\varphi^{(1)}(T), \quad P\text{-presque sûrement pour tout } \varphi \in \Phi.$$

Le problème est alors le suivant : étant donné $\varphi \rightsquigarrow I_\varphi$ continûment différentiable en moyenne quadratique de dérivée $I_\varphi^{(1)} = (I_{\varphi,i}^{(1)}; i=\overline{1,d})$ admettant une version $\tilde{I}_\varphi^{(1)} = (\tilde{I}_{\varphi,i}^{(1)}; i=\overline{1,d})$ à trajectoires continues, existe-t-il une version $(\varphi, \omega) \rightsquigarrow \tilde{I}_\varphi(\omega)$ à trajectoires dérivables de $(\varphi, \omega) \rightsquigarrow I_\varphi(\omega)$ telle que :

$$\frac{\partial}{\partial \varphi_i} \tilde{I}_\varphi(\omega) = \tilde{I}_{\varphi,i}^{(1)}(\omega), \quad \omega \in \Omega, \quad i=\overline{1,d} \quad ?$$

Supposons d'abord que $d = 1$, i.e. $\Phi \subset \mathbb{R}$: on pose alors, pour $\varphi \in \Phi$, $\omega \in \Omega$:

$$\tilde{I}_\varphi(\omega) = I_{\varphi^*}(\omega) + \int_{\varphi^*}^\varphi \tilde{I}_u^{(1)}(\omega) du$$

l'intégrale de la fonction continue $u \rightsquigarrow \tilde{I}_u^{(1)}(w)$ étant définie au sens ordinaire, φ^* étant donné dans \mathfrak{F} . Il est clair que $(w, \varphi) \rightsquigarrow \tilde{I}_\varphi(w)$ est à trajectoires continues et que :

$$\tilde{I}_\varphi^{(1)}(w) = \frac{d}{d\varphi} \tilde{I}_\varphi(w) \quad , \quad w \in \Omega .$$

De plus, pour tout $\varphi \in \mathfrak{F}$, on a :

$$I_\varphi = I_{\varphi^*} + \int_{\varphi^*}^{\varphi} \tilde{I}_u^{(1)} du$$

l'intégrale étant prise au sens de la moyenne quadratique. Alors on a aussi :

$$\tilde{I}_\varphi = I_\varphi \quad \text{P-presque sûrement, pour tout } \varphi \in \mathfrak{F} ,$$

et en définitive le couple $(\tilde{I}_\varphi, \tilde{I}_\varphi^{(1)})$ répond au problème posé.

Supposons maintenant que $d = 2$, i.e. $\varphi = (\varphi_1, \varphi_2) \in \mathfrak{F} \subset \mathbb{R}^2$.

D'après ce qui précède :

■ d'une part, pour tout φ_1 , il existe une version $(\varphi_2, w) \rightsquigarrow \tilde{I}_{\varphi_1}(\varphi_2, w)$ à trajectoires dérivables de $(\varphi_2, w) \rightsquigarrow I_{(\varphi_1, \varphi_2)}(w)$ telle que :

$$\frac{d}{d\varphi_2} \tilde{I}_{\varphi_1}(\varphi_2, w) = \tilde{I}_{(\varphi_1, \varphi_2), 2}^{(1)}(w) \quad \text{pour } w \in \Omega .$$

■ d'autre part, pour tout φ_2 , il existe une version $(\varphi_1, w) \rightsquigarrow \tilde{I}_{\varphi_2}(\varphi_1, w)$ à trajectoires dérivables de $(\varphi_1, w) \rightsquigarrow I_{(\varphi_1, \varphi_2)}(w)$ telle que :

$$\frac{d}{d\varphi_1} \tilde{I}_{\varphi_2}(\varphi_1, w) = \tilde{I}_{(\varphi_1, \varphi_2), 1}^{(1)}(w) \quad \text{pour } w \in \Omega .$$

On définit alors, pour $(\varphi_1, \varphi_2) \in \mathfrak{F}$, $w \in \Omega$:

$$\hat{I}'_{(\varphi_1, \varphi_2)}(w) = \tilde{I}_{\varphi_2}^*(\varphi_1, w) + \int_{\varphi_2^*}^{\varphi_2} \tilde{I}_{(\varphi_1, u), 2}^{(1)}(w) du$$

et

$$\hat{I}''_{(\varphi_1, \varphi_2)}(w) = \tilde{I}_{\varphi_1}^*(\varphi_2, w) + \int_{\varphi_1^*}^{\varphi_1} \tilde{I}_{(u, \varphi_2), 1}^{(1)}(w) du$$

les intégrales étant définies au sens ordinaire, $(\varphi_1^*, \varphi_2^*) = \varphi^*$ étant donné dans \mathfrak{F} . Les processus $((\varphi_1, \varphi_2), w) \rightsquigarrow \hat{I}'_{(\varphi_1, \varphi_2)}(w)$ et

$((\varphi_1, \varphi_2), w) \rightsquigarrow \hat{I}''_{(\varphi_1, \varphi_2)}(w)$ sont à trajectoires continues et tels que :

$$\hat{I}'_{(\varphi_1, \varphi_2)} = \hat{I}''_{(\varphi_1, \varphi_2)} = I_{(\varphi_1, \varphi_2)} \quad \text{P-presque sûrement, pour tout } (\varphi_1, \varphi_2) ,$$

puisque $I_{(\varphi_1, \varphi_2)}$ vérifie les mêmes équations au sens de la moyenne quadratique.

Alors P-presque sûrement, pour tout (φ_1, φ_2) on a :

$$\hat{I}'_{(\varphi_1, \varphi_2)} = \hat{I}''_{(\varphi_1, \varphi_2)} .$$

On peut donc définir $((\varphi_1, \varphi_2), \omega) \rightsquigarrow \hat{I}_{(\varphi_1, \varphi_2)}(\omega)$ à trajectoires continues tel que :

$$\hat{I}_{(\varphi_1, \varphi_2)} = I_{(\varphi_1, \varphi_2)} \quad \text{P-presque sûrement pour tout } (\varphi_1, \varphi_2)$$

$$\text{et } \frac{\partial}{\partial \varphi_1} \hat{I}_{(\varphi_1, \varphi_2)}(\omega) = \frac{\partial}{\partial \varphi_1} \hat{I}''_{(\varphi_1, \varphi_2)}(\omega) = \tilde{I}_{(\varphi_1, \varphi_2), 1}^{(1)} ,$$

$$\frac{\partial}{\partial \varphi_2} \hat{I}_{(\varphi_1, \varphi_2)}(\omega) = \frac{\partial}{\partial \varphi_2} \hat{I}'_{(\varphi_1, \varphi_2)}(\omega) = \tilde{I}_{(\varphi_1, \varphi_2), 2}^{(1)}(\omega) ; \quad \omega \in \Omega .$$

Ainsi le couple $(\hat{I}_{\varphi}, \tilde{I}_{\varphi}^{(1)})$ satisfait aux conditions imposées.

Par récurrence, il est maintenant facile de démontrer le résultat pour tout d . ■

3.2. Normalité asymptotique.

En vue d'assurer la normalité asymptotique de familles d'estimateurs du maximum de vraisemblance correspondant aux structures statistiques qui nous occuperont, nous étudions le problème suivant :

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ une base de processus, $W = (W_t ; t \geq 0)$ un processus de mouvement brownien p -dimensionnel et $\psi = (\psi_t ; t \geq 0)$ un processus mesurable à valeurs dans $\mathcal{L}(\mathbb{R}^p; \mathbb{R}^q)$ définis sur cette base, tels que :

$$P\left[\int_0^T \|\psi_t\|^2 dt < +\infty\right] = 1 , \quad \text{pour tout } T > 0 ,$$

l'intégrale stochastique $\int_0^T \psi_t dW_t$ étant alors bien définie pour tout $T > 0$.

On s'intéresse à des conditions sous lesquelles la famille de vecteurs aléatoires $\left\{ \frac{1}{\sqrt{T}} \int_0^T \psi_t dW_t ; T > 0 \right\}$ a une distribution asymptotique ($T \rightarrow +\infty$) gaussienne.

On indique d'abord un énoncé faisant la synthèse d'un résultat bien connu et d'un théorème non classique de A.F. TARASKIN [49] dont nous reproduisons la démonstration.

3.2.1. THEOREME : "Si le processus $\psi = (\psi_t ; t \geq 0)$ est non aléatoire i.e. ψ_t ne dépend pas de $\omega \in \Omega$, alors, pour tout $T > 0$, l'intégrale stochastique $\int_0^T \psi_t dW_t$ est un vecteur aléatoire de loi gaussienne $N(0, \int_0^T \psi_s \cdot {}^t \psi_s ds)$. Si le processus aléatoire $\psi = (\psi_t ; t \geq 0)$ satisfait aux conditions :

$$(H 1) \quad P\left\{ \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \psi_s \cdot {}^t \psi_s ds = \Lambda \right\} = 1$$

$$(H 2) \quad \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T E(\psi_s \cdot {}^t \psi_s) ds = \Lambda$$

alors la famille de vecteurs aléatoires $\left\{ \frac{1}{\sqrt{T}} \int_0^T \psi_t dW_t ; T > 0 \right\}$ est asymptotiquement gaussienne $N(0, \Lambda)$."

Démonstration : La première partie de l'énoncé est une conséquence facile de la définition de l'intégrale stochastique (voir par exemple [4], Corollaire 4.5.6). Pour démontrer la seconde partie du théorème, on montre d'abord que pour tout $\lambda \in \mathbb{R}^q$, la famille $\left\{ \langle \lambda, \frac{1}{\sqrt{T}} \int_0^T \psi_s dW_s \rangle ; T > 0 \right\}$ de variables aléatoires converge en loi. On note

$$U_T^\lambda = \frac{1}{\sqrt{T}} \int_0^T {}^t \lambda \psi_s dW_s ; \quad T > 0 .$$

Si λ est tel que ${}^t \lambda \wedge \lambda > 0$ on a, compte-tenu de (H 1), avec probabilité 1 :

$$(14) \quad \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T {}^t \lambda \psi_s \cdot {}^t \psi_s \cdot \lambda ds = {}^t \lambda \wedge \lambda \neq 0$$

et donc, avec probabilité 1 :

$$\lim_{T \rightarrow +\infty} \int_0^T {}^t \lambda \psi_s \cdot {}^t \psi_s \cdot \lambda ds = +\infty .$$

On introduit alors le temps d'arrêt :

$$\tau_T^\lambda = \inf\{u \geq 0 : \int_0^u t_{\lambda \psi_s} \cdot t_{\psi_s} \cdot \lambda ds \geq T \cdot t_{\lambda \wedge \lambda}\}$$

et

$$V_T^\lambda = (t_{\lambda \wedge \lambda})^{-1/2} \cdot \int_0^{\tau_T^\lambda} t_{\lambda \psi_s} \cdot t_{\psi_s} dW_s .$$

D'après un résultat classique sur le changement de temps dans une martingale (voir par exemple [43]), le processus $(V_T^\lambda ; T \geq 0)$ est un mouvement brownien standard.

On évalue alors :

$$\begin{aligned} & E\{ |(t_{\lambda \wedge \lambda})^{-1/2} \cdot U_T^\lambda - T^{-1/2} V_T^\lambda|^2 \} \\ &= (T \cdot t_{\lambda \wedge \lambda})^{-1} E\{ \left| \int_0^T t_{\lambda \psi_s} dW_s - \int_0^{\tau_T^\lambda} t_{\lambda \psi_s} dW_s \right|^2 \} \\ (15) \quad &= (T \cdot t_{\lambda \wedge \lambda})^{-1} E\{ \int_{T \wedge \tau_T^\lambda}^{T \vee \tau_T^\lambda} t_{\lambda \psi_s} \cdot t_{\psi_s} \cdot \lambda ds \} \\ &= (t_{\lambda \wedge \lambda})^{-1} \cdot E\{ \left| \frac{1}{T} \int_0^T t_{\lambda \psi_s} \cdot t_{\psi_s} \cdot \lambda ds - \frac{1}{T} \int_0^{\tau_T^\lambda} t_{\lambda \psi_s} \cdot t_{\psi_s} \cdot \lambda ds \right| \} \\ &= (t_{\lambda \wedge \lambda})^{-1} \cdot E\{ \left| \frac{1}{T} \int_0^T t_{\lambda \psi_s} \cdot t_{\psi_s} \cdot \lambda ds - t_{\lambda \wedge \lambda} \right| \} . \end{aligned}$$

On démontre que le dernier membre de la suite d'égalités (15) tend vers 0 lorsque $T \rightarrow +\infty$. Or comme d'une part, d'après (14), l'expression se trouvant sous le symbole espérance mathématique tend en probabilité vers 0 et, que d'autre part, d'après (H2),

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T E(t_{\lambda \psi_s} \cdot t_{\psi_s} \cdot \lambda) ds = t_{\lambda \wedge \lambda} ,$$

on a aussi

$$\lim_{T \rightarrow +\infty} \frac{1}{T} E \int_0^T t_{\lambda \psi_s} \cdot t_{\psi_s} \lambda ds = t_{\lambda \wedge \lambda}$$

puis

$$\lim_{T \rightarrow +\infty} E\{ \left| \frac{1}{T} \int_0^T t_{\lambda \psi_s} \cdot t_{\psi_s} \lambda ds - t_{\lambda \wedge \lambda} \right| \} = 0 .$$

Ainsi, d'après (15), $(t_{\lambda \wedge \lambda})^{-1/2} \cdot U_T^\lambda - T^{-1/2} V_T^\lambda$ converge en moyenne quadratique vers 0 lorsque $T \rightarrow +\infty$. La famille $\{U_T^\lambda ; T > 0\}$ est donc asymptotiquement gaussienne $N(0, t_{\lambda \wedge \lambda})$.

Si maintenant λ est tel que ${}^t_{\lambda\Lambda\lambda} = 0$, on a :

$$\lim_{T \rightarrow +\infty} \frac{1}{T} E \int_0^T {}^t_{\lambda} \psi_s \cdot {}^t_{\psi_s} \lambda ds = 0$$

et donc :

$$\lim_{T \rightarrow +\infty} E |U_T^\lambda|^2 = 0.$$

En définitive, on a, pour tout $\lambda \in \mathbb{R}^q$, et tout $u \in \mathbb{R}$:

$$\lim_{T \rightarrow +\infty} E(e^{iuU_T^\lambda}) = e^{-\frac{u^2}{2} {}^t_{\lambda\Lambda\lambda}}$$

ce qui signifie que la fonction caractéristique du vecteur aléatoire $\frac{1}{\sqrt{T}} \int_0^T \psi_s dW_s$ converge ponctuellement ($T \rightarrow +\infty$) vers la fonction caractéristique d'une loi $N(0, \Lambda)$. ■

Dans [50] A.F. TARASKIN étudie la situation plus générale où $(W_t; t \geq 0)$ est une martingale continue de carré intégrable, la fonction normative pouvant être autre que $\frac{1}{\sqrt{T}}$.

Nous indiquons maintenant un résultat tel qu'il a été annoncé dans A. LE BRETON [30] et démontré dans A. LE BRETON [31] à un stade de notre travail où nous n'avions pas connaissance des articles [49] et [50]. Dans ce paragraphe nous le présentons comme corollaire du théorème 3.2.1. Dans le chapitre suivant nous en donnerons une démonstration directe.

3.2.2. PROPOSITION : "Soit θ une matrice $q \times q$ dont les valeurs propres ont toutes des parties réelles strictement négatives et G une matrice élément de $\mathfrak{L}(\mathbb{R}^p; \mathbb{R}^q)$. Soit $(Z_t)_{t \geq 0}$ un processus centré gaussien (strictement) stationnaire solution forte de l'équation différentielle stochastique :

$$(16) \quad dZ_t = \theta Z_t dt + G dW_t \quad ; \quad t \geq 0 .$$

Alors la famille de matrices aléatoires $\left\{ \frac{1}{\sqrt{T}} \int_0^T dW_s \cdot {}^t Z_s ; T > 0 \right\}$ converge en loi vers une matrice aléatoire M de loi gaussienne $N(0, \Gamma)$ où :

$$E(M_{ij} M_{k\ell}) = \Gamma_{ijk\ell} = \delta_{ik} \cdot (E(Z_0 \cdot {}^t Z_0))_{j\ell} ; \quad i, k \in \{1, \dots, p\}, \quad j, \ell \in \{1, \dots, q\}$$

Démonstration : Remarquons d'abord que θ et G étant données, vérifiant les hypothèses, l'équation matricielle :

$$(17) \quad \theta Q + Q {}^t \theta = -G {}^t G$$

admet une solution Q qui est une matrice de covariance $q \times q$ (voir par exemple E.B. LEE - L. MARKUS [35], p. 198).

D'autre part toute solution forte de (18) peut s'écrire :

$$Z_t = e^{t\theta} Z_0 + \int_0^t e^{(t-s)\theta} G dW_s \quad ; \quad t \geq 0$$

et si Z_0 est un vecteur aléatoire centré de matrice de covariances Q donnée par (19), alors le processus $(Z_t)_{t \geq 0}$ est stationnaire au sens large, de moyenne nulle et de fonction de covariance $R(h) = e^{h\theta} \cdot Q$ si $h \geq 0$. En particulier si Z_0 est de loi gaussienne le processus $(Z_t)_{t \geq 0}$ est gaussien stationnaire (voir par exemple [4] Théorème 8.2.12). Alors ce processus admet la densité spectrale :

$$f(\lambda) = \frac{1}{2\pi} (\theta - i\lambda \mathbb{1}_q)^{-1} \cdot (G \cdot {}^t G) \cdot ({}^t \theta + i\lambda \mathbb{1}_q)^{-1}$$

$\mathbb{1}_q$ désignant la matrice unité $q \times q$. Un théorème ergodique (voir [45] Yu. A. ROZANOV, pp. 157 et 163) assure que dans ces conditions, on a :

$$(18) \quad \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T Z_s \cdot {}^t Z_s ds = R(0) = Q = E(Z_0 \cdot {}^t Z_0) ,$$

au sens de la convergence P presque sûre.

Comme le processus est stationnaire on a aussi, pour tout $T > 0$,

$$\frac{1}{T} \int_0^T E(Z_s \cdot {}^t Z_s) ds = E(Z_0 \cdot {}^t Z_0)$$

CHAPITRE II

ESTIMATION DANS LES MODELES DIFFERENTIELS STOCHASTIQUES A COEFFICIENTS DEPENDANT LINEAIREMENT DES PARAMETRES INCONNUS

Un exemple simple de modèle de la théorie du contrôle optimal stochastique est celui défini par l'équation d'évolution stochastique :

$$\begin{cases} dX_t = \theta X_t dt + u_t dt + GdW_t & ; \quad t \geq 0 \\ X_0 = X(0) \end{cases} \quad \theta \text{ et } G \text{ matrices données,}$$

où si X_t , vecteur d'état à l'instant t du système à contrôler, est supposé complètement observable, le contrôle u_t à l'instant t peut dépendre des observations faites sur le système jusqu'à cet instant. La stratégie de contrôle est à choisir dans une certaine classe de stratégies admissibles (par exemple, celle des stratégies $u_{(.)}$ de la forme $u_t = KX_t$ où K est une matrice de dimension adéquate) de manière à minimiser un certain critère, par exemple :

$$E \int_0^{T_0} L(X_t, u_t, t) dt \quad , \quad T_0 \text{ donné, } L \text{ fonction donnée.}$$

Bien entendu l'étude du problème de contrôle nécessite la connaissance des coefficients θ et G de l'équation d'évolution, la réponse au problème d'optimisation dépendant de ces coefficients. Dans de nombreuses situations ces paramètres sont sinon inconnus du moins imparfaitement connus. Il se pose alors simultanément un problème statistique : celui de leur estimation. Deux cas peuvent se présenter :

1er cas : L'observateur (le futur contrôleur) a la possibilité d'expérimenter, éventuellement de façon répétitive. Il met alors en action

une fonction de contrôle choisie a priori (par exemple $u_t \equiv 0$ si la stratégie correspondante est admissible) et observe une trajectoire du processus d'état du système. Répétant l'expérience il dispose d'un échantillon au vu duquel il envisage l'estimation des paramètres. Il étudie ensuite le problème de contrôle optimal pour les valeurs estimées et contrôle enfin le système en conséquence.

2e cas : L'observateur (en même temps contrôleur) doit contrôler le système en temps réel ce qui exclut la possibilité de répéter l'expérience fournissant l'observation sur laquelle est basée l'estimation. Il semble alors raisonnable que sa démarche consiste d'abord à étudier le problème de contrôle sous les diverses hypothèses possibles sur les paramètres ainsi que celui d'estimation de ces paramètres au vu d'une seule observation. Puis il lance le système avec un contrôle choisi a priori et s'efforce d'estimer les vraies valeurs des coefficients en temps réel et d'adapter son contrôle à cette estimation compte-tenu de son étude préalable.

Nous nous plaçons ici dans la perspective de la deuxième situation en étudiant le problème d'estimation de paramètres dans un modèle différentiel stochastique sans contrôle, au vu de l'observation d'une seule trajectoire du processus qu'il engendre. La première partie de ce chapitre traite du modèle linéaire dans le cas où l'observation a lieu en temps continu. Dans une seconde partie nous étudions le cas où l'observation a lieu en temps discret. Enfin nous envisageons quelques extensions et applications de cette étude.

§ 1. - ESTIMATION DANS LE MODELE LINEAIRE. CAS DE L'OBSERVATION EN TEMPS CONTINU.

Nous pourrions, compte-tenu de la présentation du modèle général faite dans le chapitre précédent, nous contenter de dire que l'objet de l'étude est l'estimation du paramètre (θ, G) dans la structure statistique :

$$(C, S_T, \{P_{\theta, G}^T; (\theta, G) \in \Theta \times \mathcal{G}\})$$

où $\Theta \subset \mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$ et $\mathcal{G} \subset \mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$

et, pour tout $(\theta, G) \in \Theta \times \mathcal{G}$,

$$\frac{dP_{\theta, G}^T}{dP_{0, G}^T} = \exp\left\{\int_0^T \langle \theta \Pi_s, (G^t G)^{-1} d\Pi_s \rangle - \frac{1}{2} \int_0^T \langle \theta \Pi_s, (G^t G)^{-1} \theta \Pi_s \rangle ds\right\}$$

où l'intégrale stochastique est définie relativement à la mesure de Wiener de covariance $G^t G$ notée $P_{0, G}^T$.

Toutefois nous choisissons dans ce contexte simple de donner une présentation directe du problème et d'assurer, à l'aide d'énoncés du chapitre I, § 1, qu'on est conduit à le traduire dans les termes précédents. Le contenu de ce paragraphe a été annoncé dans A. LE BRETON [30] et développé dans [31].

1.1. Formulation directe du problème. Structure statistique.

On observe sur un intervalle de temps $[0, T]$, $T > 0$ une fonction continue $x(\cdot) = (x_t; t \geq 0) \in C$, trajectoire d'état d'un système dynamique stochastique.

Θ (resp. \mathcal{G}) étant un sous-ensemble de $\mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$ (resp. $\mathcal{L}^*(\mathbb{R}^p; \mathbb{R}^p)$ ensemble des matrices $p \times p$ inversibles), on suppose qu'il existe un couple (θ^*, G^*) inconnu dans $\Theta \times \mathcal{G}$, dit "vraie valeur de (θ, G) ", tel que pour $(\theta, G) = (\theta^*, G^*)$, $x(\cdot)$ soit une trajectoire d'un processus $(X_t; t \geq 0)$ solution de l'équation différentielle stochastique :

$$(1) \quad dX_t = \theta X_t dt + G dW_t ; \quad t \geq 0, \quad X_0 = 0$$

où $(W_t)_{t \geq 0}$ est un processus de mouvement brownien standard, i.e. il existe un espace probabilisé (Ω, \mathcal{A}, P) muni d'une famille croissante $(\mathcal{A}_t)_{t \geq 0}$ de sous-tribus de \mathcal{A} , un mouvement brownien p -dimensionnel $(W_t)_{t \geq 0}$ relativement à $(\mathcal{A}_t)_{t \geq 0}$ et une fonction aléatoire $(X_t)_{t \geq 0}$ continue et adaptée à $(\mathcal{A}_t)_{t \geq 0}$, telle que :

$$X_t = \int_0^t \theta X_s ds + GW_t ; t \geq 0$$

et que, pour un élément $\omega \in \Omega$, on ait $x(\cdot) = (X_t(\omega); t \geq 0)$.

Pour $(\theta, G) \in \Theta \times \mathcal{G}$ et $(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t \geq 0}, P, (W_t)_{t \geq 0})$ donnés on est assuré (voir chapitre I, théorème 1.3.1.) de l'existence et de l'unicité de la solution forte de l'équation (1) et donc de l'unicité en loi au sens que toutes les solutions :

$$(\Omega, \mathcal{A}, (\mathcal{A}_t)_{t \geq 0}, P, (W_t)_{t \geq 0}, (X_t)_{t \geq 0})$$

induisent la même mesure sur l'espace des trajectoires C . Cette loi qui ne dépend d'ailleurs que de $(\theta, G^t G)$ peut ainsi être définie comme étant la mesure $\mu_{\theta, G}$ induite sur C par la solution particulière suivante :

$$(C, \mathcal{S}, (\mathcal{S}_t)_{t \geq 0}, \mu_*, (\Pi_t)_{t \geq 0}, (Z_t)_{t \geq 0})$$

où $(Z_t)_{t \geq 0}$ est la solution forte de l'équation :

$$dZ_t = \theta Z_t dt + G d\Pi_t ; t \geq 0, Z_0 = 0$$

relativement à l'espace probabilisé $(C, \mathcal{S}, (\mathcal{S}_t)_{t \geq 0}, \mu_*)$ i.e.

$$Z_t = \int_0^t e^{(t-s)\theta} G d\Pi_s ; t \geq 0$$

qui est un processus gaussien centré de fonction de covariance :

$$r(s, t) = e^{s\theta} \left\{ \int_0^{s \wedge t} e^{-u\theta} G^t G e^{-u\theta} du \right\} e^{t\theta}.$$

On est alors amené naturellement à poser le problème d'estimation du paramètre (θ, G) , au vu de l'observation de $x(\cdot)$ sur l'intervalle de temps $[0, T]$, dans la structure statistique :

$$(C, \mathcal{S}_T, \{\mu_{\theta, G/S_T}; (\theta, G) \in \Theta \times \mathcal{G}\}) .$$

1.1.1. LEMME : "Pour tout $(\theta, G) \in \Theta \times \mathcal{G}$ et tout $T > 0$ les mesures $\mu_{\theta, G/S_T}$ et $\mu_{0, G/S_T}$ sont équivalentes, la dérivée de Radon-Nikodym étant donnée par :

$$(2) \quad \frac{d\mu_{\theta, G}}{d\mu_{0, G}} / S_T = \exp\left\{\int_0^T \langle \theta \Pi_t, (G^t G)^{-1} d\Pi_t \rangle - \frac{1}{2} \int_0^T \langle \theta \Pi_t, (G^t G)^{-1} \theta \Pi_t \rangle dt\right\}$$

où l'intégrale stochastique est définie relativement à la base de processus $(C, S, (S_t)_{t \geq 0}, \mu_{0, G})$.

Démonstration : Considérons le processus $(Z_t^1; t \geq 0)$ défini relativement à $(C, S, (S_t)_{t \geq 0}, \mu_*)$ par :

$$dZ_t^1 = G d\Pi_t \quad ; \quad t \geq 0 \quad ; \quad Z_0^1 = 0 \quad , \quad \text{i.e.} \quad Z_t^1 = G \Pi_t \quad , \quad t \geq 0 \quad ,$$

et la mesure $\tilde{\mu}^T$ sur S_T , absolument continue par rapport à μ_*^T avec pour dérivée de Radon-Nikodym

$$\frac{d\tilde{\mu}^T}{d\mu_*^T} = \exp\left\{\int_0^T \langle G^{-1} \theta Z_t^1, d\Pi_t \rangle - \frac{1}{2} \int_0^T \langle G^{-1} \theta Z_t^1, G^{-1} \theta Z_t^1 \rangle dt\right\} .$$

Les théorèmes I, 1.3.3. et 1.2.1. assurent que $\tilde{\mu}^T$ est une probabilité et que de plus le processus $\{\tilde{\Pi}_t; t \in [0, T]\}$ défini par :

$$\tilde{\Pi}_t = \Pi_t - \int_0^t G^{-1} \theta Z_s^1 ds \quad ; \quad t \in [0, T]$$

est un processus de Wiener sur la base de processus $(C, (S_t)_{0 \leq t \leq T}, \tilde{\mu}^T)$ et $Z_t^1 = \int_0^t \theta Z_s^1 ds + G \tilde{\Pi}_t \quad ; \quad t \in [0, T]$.

D'après le théorème I, 1.2.3. la loi du processus $(Z_t^1; t \in [0, T]; \tilde{\mu}^T)$ est équivalente à celle du processus $(Z_t^1; t \in [0, T]; \mu_*^T)$, la dérivée de Radon-Nikodym s'écrivant :

$$\exp\left\{\int_0^T \langle G^{-1} \theta \Pi_t, G^{-1} d\Pi_t \rangle - \frac{1}{2} \int_0^T \langle G^{-1} \theta \Pi_t, G^{-1} \theta \Pi_t \rangle dt\right\}$$

où l'intégrale stochastique est définie relativement à la base $(C, S, (S_t)_{0 \leq t \leq T}, \mu_{0, G} / S_T)$. Le résultat du lemme est alors acquis compte-tenu de l'unicité (en loi) de la solution de (1). ■

Ainsi le problème d'estimation du paramètre (θ, G) ou plus exactement du paramètre $(\theta, G^t G)$ se traduit bien dans les termes envisagés au début du paragraphe 1.

1.2. Estimation de la matrice $G^t G$.

On considère la suite de statistiques $\{\hat{\sigma}_t^N ; N \geq 1\}$ définie sur la structure statistique concernée par :

$$(3) \quad \begin{aligned} \hat{\sigma}_t^N(x) &= \frac{1}{t} \sum_{i=1}^{2^N} \left(x_{i \frac{t}{2^N}} - x_{(i-1) \frac{t}{2^N}} \right) \cdot \left(x_{i \frac{t}{2^N}} - x_{(i-1) \frac{t}{2^N}} \right) \\ &= \frac{1}{t} \sum_{i=1}^{2^N} \left(x_{i \frac{t}{2^N}} - x_{(i-1) \frac{t}{2^N}} \right)^2 \otimes \end{aligned}$$

pour $t \in]0, T]$ et $x = (x_t ; t \geq 0) \in C$.

On a alors la proposition suivante, dans les termes de la définition I, 1.4.1. :

1.2.1. PROPOSITION : "Dans la paramétrisation $\Theta \times \mathcal{G}$ le paramètre $G^t G$ est fortement identifiable presque sûrement sur tout intervalle de temps $[0, t]$, $t > 0$, par la suite $\{\hat{\sigma}_t^N ; N \geq 1\}$ de statistiques définies par (3) i.e. pour tout $(\theta, G) \in \Theta \times \mathcal{G}$ on a, pour tout $t \in]0, T]$:

$$(4) \quad \lim_{N \rightarrow +\infty} \hat{\sigma}_t^N = G^t G \quad P_{\theta, G}^T \text{ presque sûrement. "}$$

Cette proposition est une conséquence immédiate du lemme classique suivant :

1.2.2. LEMME : "Soit $(W_t ; t \geq 0)$ un processus de mouvement brownien standard p-dimensionnel défini sur $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ et $0 \leq s < t$. Soit, pour tout $N \geq 1$, $s = t_0^N < t_1^N < \dots < t_{K_N}^N = t$ une subdivision de l'intervalle $[s, t]$ telle que $\delta_N = \max_{k=1, K_N} |t_k^N - t_{k-1}^N|$. Alors

$$(5) \quad \lim_{\delta_N \rightarrow 0} m.p. \sum_{k=1}^{K_N} (W_{t_k^N} - W_{t_{k-1}^N})^2 \otimes = (t-s) \mathbf{1}_p .$$

De plus, si $\sum_{N=1}^{\infty} \delta_N < +\infty$, alors la convergence dans (5) a lieu avec probabilité 1."

Démonstration [4] : Soit ${}^tW_s = (W_s^1, W_s^2, \dots, W_s^p)$. Si

$$S_N^{ij} = \sum_{k=1}^{K_N} (W_{t_k}^i - W_{t_{k-1}}^i) (W_{t_k}^j - W_{t_{k-1}}^j) ,$$

alors :

$$E(S_N^{ij}) = (t-s) \delta_{ij}$$

$$\begin{aligned} \text{et } \sigma^2(S_N^{ij}) &= \sum_{k=1}^{K_N} (E(W_{t_k}^i - W_{t_{k-1}}^i)^2 (W_{t_k}^j - W_{t_{k-1}}^j)^2 - \delta_{ij} (t_k^N - t_{k-1}^N)^2) \\ &= (1 + \delta_{ij}) \sum_{k=1}^{K_N} (t_k^N - t_{k-1}^N)^2 \\ &\leq 2(t-s) \delta_N . \end{aligned}$$

Ainsi $\lim_{\delta_N \rightarrow 0} E |S_N^{ij} - \delta_{ij}(t-s)|^2 = 0$ ce qui démontre (5).

Si $\sum_{N=1}^{\infty} \delta_N < +\infty$, alors $\sum_{N=1}^{\infty} E |S_N^{ij} - \delta_{ij}(t-s)|^2 < +\infty$, ce qui suffit à assurer la convergence presque sûre dans (5). ■

1.2.3. Démonstration de la proposition 1.2.1.

Le processus $(G^{-1}\Pi_t ; t \in [0, T])$ est un processus de mouvement brownien standard relativement à la mesure $P_{0,G}^T$. D'après le lemme 1.2.2. on a donc :

$$\lim_{N \rightarrow +\infty} \frac{1}{t} \sum_{i=1}^{2N} (G^{-1}\Pi_{i \frac{t}{2N}} - G^{-1}\Pi_{(i-1) \frac{t}{2N}})^2 = \mathbb{1}_P$$

au sens de la convergence $P_{0,G}^T$ presque sûre, i.e.

$$\lim_{N \rightarrow +\infty} G^{-1}\hat{\sigma}_t^N ({}^tG)^{-1} = \mathbb{1}_P$$

et par suite,

$$\lim_{N \rightarrow +\infty} \hat{\sigma}_t^N = G {}^tG , \quad P_{0,G}^T \text{ presque sûrement.}$$

Comme d'après le lemme 1.1.1. pour tout $\theta \in \Theta$ les mesures $P_{0,G}^T$ et $P_{\theta,G}^T$ sont équivalentes, on a aussi (4). ■

1.2.4. Remarques : a) Dans le §2 de ce chapitre, dans le lemme 2.2.3. nous précisons la qualité de l'estimateur $\hat{\sigma}_t^N$: la suite de matrices aléatoires $\{\delta_N^{-1/2} (\hat{\sigma}_t^N - G {}^tG) ; N \geq 1\}$ est bornée en probabilité.

b) Ainsi la matrice $G^t G$ peut être calculée, avec probabilité 1, à l'aide de l'observation d'une trajectoire du processus, sur tout intervalle de temps fini. Partant il est légitime, pour étudier le problème d'estimation de θ de supposer que la matrice $G^t G$ est complètement connue. C'est ce que nous ferons dans ce qui suit en posant $G^t G = \Gamma$.

1.3. Estimation de la matrice θ

La structure statistique du problème est maintenant :

$$(C, S_T, \{P_{\theta, \Gamma}^T; \theta \in \Theta\})$$

où $P_{\theta, \Gamma}^T = \mu_{\theta, \Gamma} / S_T$, avec $\mu_{\theta, \Gamma} = \mu_{\theta, G}$.

D'après le lemme 1.1.1. cette structure est dominée par la mesure $P_{0, \Gamma}^T$ (complètement définie par Γ), la fonction de vraisemblance étant donnée par :

$$(6) \quad \frac{dP_{\theta, \Gamma}^T}{dP_{0, \Gamma}^T} = \exp\left\{\int_0^T \langle \theta \Pi_t, \Gamma^{-1} d\Pi_t \rangle - \frac{1}{2} \int_0^T \langle \theta \Pi_t, \Gamma^{-1} \theta \Pi_t \rangle dt\right\}.$$

Remarquons que la structure est de type exponentiel, admettant la statistique exhaustive :

$$\left(\int_0^T d\Pi_t \cdot {}^t \Pi_t, \int_0^T \Pi_t \cdot {}^t \Pi_t dt\right)$$

où l'intégrale stochastique est définie relativement à $P_{0, \Gamma}^T$. En effet, on peut écrire (6) sous la forme :

$$(7) \quad \frac{dP_{\theta, \Gamma}^T}{dP_{0, \Gamma}^T} = \exp\left\{\text{tr}\left[\Gamma^{-1} \theta \int_0^T \Pi_t \cdot d\Pi_t\right] - \frac{1}{2} \text{tr}\left[{}^t \theta \Gamma^{-1} \theta \int_0^T \Pi_t \cdot {}^t \Pi_t dt\right]\right\}.$$

En particulier dans le cas scalaire, i.e. $p = 1$, la statistique $(\Pi_T^2 - \Gamma T, \int_0^T \Pi_t^2 dt)$ est exhaustive. La fonction de Log-vraisemblance sur la structure est donnée par :

$$(8) \quad L_{\Gamma, T}(\theta, \cdot) = \int_0^T \langle \theta \Pi_t, \Gamma^{-1} d\Pi_t \rangle - \frac{1}{2} \int_0^T \langle \theta \Pi_t, \Gamma^{-1} \theta \Pi_t \rangle dt$$

i. e.

$$(8)' \quad L_{\Gamma, T}(\theta, x) = \left(\int_0^T \langle \theta \Pi_t, \Gamma^{-1} d\Pi_t \rangle \right)(x) - \frac{1}{2} \int_0^T \langle \theta x_t, \Gamma^{-1} \theta x_t \rangle dt$$

pour $x \in C$, où l'intégrale stochastique est définie relativement à $P_{0, \Gamma}^T$.
Compte-tenu de la forme particulière de (8), le choix d'une version régulière de la fonction de Log-vraisemblance consiste en le seul choix d'une version de l'intégrale stochastique.

Si $\theta = ((\theta_{ij}))_{i,j=\overline{1,p}}$, $\Gamma^{-1} = ((\rho_{ij}))_{i,j=\overline{1,p}}$ et ${}^t\Pi_t = (\Pi_t^1, \Pi_t^2, \dots, \Pi_t^p)$ on peut écrire :

$$L_{\Gamma, T}(\theta, \cdot) = \int_0^T \sum_{i=1}^p \left(\sum_{j=1}^p \theta_{ij} \Pi_s^j \right) \left(\sum_{j=1}^p \rho_{ij} d\Pi_s^j \right) - \frac{1}{2} \int_0^T \left(\sum_{i=1}^p \left(\sum_{j=1}^p \theta_{ij} \Pi_s^j \right) \left(\sum_{\ell=1}^p \rho_{i\ell} \left(\sum_{r=1}^p \theta_{er} \Pi_s^r \right) \right) \right) ds .$$

Le système des équations de maximum de vraisemblance s'écrit alors :

$$\frac{\partial}{\partial \theta_{uv}} L_{\Gamma, T}(\theta, \cdot) = 0 \quad ; \quad u = \overline{1,p}, \quad v = \overline{1,p}$$

$$\text{soit} \quad \int_0^T \left(\sum_{\ell=1}^p \rho_{u\ell} d\Pi_s^\ell \right) \Pi_s^v - \int_0^T \left(\sum_{\ell=1}^p \rho_{u\ell} \left(\sum_{r=1}^p \theta_{\ell r} \Pi_s^r \right) \right) \Pi_s^v ds = 0 \quad ; \quad u = \overline{1,p}, \quad v = \overline{1,p} .$$

Ainsi la matrice :

$$\frac{d}{d\theta} L_{\Gamma, T}(\theta, \cdot) = \left(\left(\frac{\partial}{\partial \theta_{u,v}} L_{\Gamma, T}(\theta, \cdot) \right) \right)_{u,v=\overline{1,p}}$$

n'est autre que :

$$\int_0^T \Gamma^{-1} d\Pi_s \cdot {}^t\Pi_s - \int_0^T \Gamma^{-1} \theta \Pi_s \cdot {}^t\Pi_s ds ,$$

et le système des équations de vraisemblance s'écrit donc :

$$\int_0^T d\Pi_s \cdot {}^t\Pi_s = \theta \int_0^T \Pi_s \cdot {}^t\Pi_s ds .$$

Nous sommes alors en mesure d'énoncer la proposition suivante :

1.3.1. PROPOSITION : "L'estimateur de maximum de vraisemblance de θ au vu de l'observation sur $[0, T]$ est donné par :

$$(9) \quad \hat{\theta}_{\Gamma, T} = \left[\int_0^T d\Pi_s \cdot {}^t\Pi_s \right] \left[\int_0^T \Pi_s \cdot {}^t\Pi_s ds \right]^{-1}$$

où l'intégrale stochastique est définie relativement à $P_{0,\Gamma}^T$."

Démonstration : Elle consiste simplement à vérifier que la matrice aléatoire symétrique semi-définie positive $\int_0^T \Pi_s \cdot {}^t \Pi_s ds$ est $P_{0,\Gamma}^T$ presque sûrement non singulière. Il est clair qu'il suffit de considérer le cas où $\Gamma = \mathbb{1}_p$ i.e. $P_{0,\Gamma}^T = \mu_*^T$. Définissons :

$$\begin{aligned} A_0 &= \{x \in C : \exists h \in \mathbb{R}^p, h \neq 0 / \langle [\int_0^T x_s \cdot {}^t x_s ds] h, h \rangle = 0 \} \\ &= \{x \in C : \exists h \in \mathbb{R}^p, h \neq 0 / \int_0^T \langle x_s, h \rangle^2 ds = 0 \} \\ &= \{x \in C : \exists h \in \mathbb{R}^p, h \neq 0 / \forall s \in [0, T], \langle x_s, h \rangle = 0 \} . \end{aligned}$$

Pour $\{s_1, \dots, s_{p+1}\} \subset [0, T]$, $0 \leq s_1 < s_2 < \dots < s_{p+1} \leq T$, si $x \in A_0$ il existe $h \in \mathbb{R}^p$, $h \neq 0$, tel que :

$$\langle x_{s_{i+1}} - x_{s_i}, h \rangle = 0 \text{ pour } i = \overline{1, p} ,$$

et donc aussi

$$\det \{ x_{s_2} - x_{s_1}, \dots, x_{s_{p+1}} - x_{s_p} \} = 0 .$$

Ainsi $A_0 \subset A_0^* = \{x \in C : \det \{ x_{s_2} - x_{s_1}, \dots, x_{s_{p+1}} - x_{s_p} \} = 0 \}$.

L'application D de $(\mathbb{R}^p)^p$ dans \mathbb{R} définie par :

$$D(v_1, \dots, v_p) = \det \{ v_1, \dots, v_p \} , \quad \forall v_i \in \mathbb{R}^p, i = \overline{1, p}$$

étant continue, le sous-ensemble Δ de $(\mathbb{R}^p)^p$

$$\Delta = \{ (v_1, \dots, v_p) \in (\mathbb{R}^p)^p / \det \{ v_1, \dots, v_p \} = 0 \}$$

est fermé. Par suite A_0^* est mesurable dans (C, S_T) et on a :

$$\mu_*^T(A_0^*) = P_\tau(\Delta)$$

où P_τ est la loi de probabilité de l'élément aléatoire

$\tau = (\Pi_{s_2} - \Pi_{s_1}, \dots, \Pi_{s_{p+1}} - \Pi_{s_p})$ défini sur (C, S_T, μ_*^T) . P_τ étant une

mesure gaussienne non singulière sur $(\mathbb{R}^p)^p$, et Δ étant de mesure de Lebesgue nulle dans $(\mathbb{R}^p)^p$, le résultat du lemme est démontré. ■

1.3.2. Remarques.

a) On peut voir aisément que

$$\frac{\partial^2}{\partial \theta_{k\ell} \partial \theta_{ij}} L_{\Gamma, T}(\theta, \cdot) = -(\Gamma^{-1})_{i,k} \left(\int_0^T \Pi_s \cdot {}^t \Pi_s ds \right)_{\ell, j} ; \quad k, \ell, i, j = \overline{1, p} .$$

Comme la matrice $\int_0^T \Pi_s \cdot {}^t \Pi_s ds$ est $P_{0, \Gamma}^T$ presque sûrement définie positive, la matrice $\left(\frac{\partial}{\partial \theta_{k\ell} \partial \theta_{ij}} L_{\Gamma, T}(\theta, \cdot) \right)_{k, \ell, i, j = \overline{1, p}}$ est $P_{0, \Gamma}^T$ presque sûrement définie négative. $\hat{\theta}_{\Gamma, T}$ réalise donc, $P_{0, \Gamma}^T$ presque sûrement, le maximum de $L_{\Gamma, T}(\theta, \cdot)$, $\theta \in \mathfrak{L}(\mathbb{R}^p; \mathbb{R}^p)$.

b) Formellement, on peut écrire :

$$L_{\Gamma, T}(\theta, x) = -\frac{1}{2} \left[\int_0^T \frac{\|dx_s - \theta x_s ds\|^2}{ds} - \int_0^T \frac{\|dx_s\|^2}{ds} \right]$$

et, pour maximiser la fonction de vraisemblance, il s'agit de minimiser

$$\int_0^T \frac{\|dx_s - \theta x_s ds\|^2}{ds}$$

expression qui apparaît sous la forme d'une "somme de carrés de résidus".

L'estimateur $\hat{\theta}_{\Gamma, T}$ est donc aussi un estimateur de "moindres carrés".

c) Dans le cas où $p = 1$, on a, par la formule de Itô :

$$\begin{aligned} \theta_{\Gamma, T}(x) &= \left(\int_0^T \Pi_s d\Pi_s \right)(x) / \left(\int_0^T \Pi_s^2 ds \right)(x) \\ &= \frac{\Pi_T^2(x) - \Gamma T}{2 \int_0^T \Pi_s^2(x) ds} \\ &= \frac{x_T^2 - \Gamma T}{2 \int_0^T x_s^2 ds} . \end{aligned}$$

d) Dans le cas où $p \geq 2$ et où Θ est le sous-ensemble de $\mathfrak{L}(\mathbb{R}^p; \mathbb{R}^p)$ constitué par les matrices symétriques, on peut montrer aisément que l'estimateur de maximum de vraisemblance de θ est donné par la solution de l'équation

$$\int_0^T d\Pi_s \cdot {}^t \Pi_s + \int_0^T \Pi_s \cdot {}^t d\Pi_s - \left[\theta \int_0^T \Pi_s \cdot {}^t \Pi_s ds + \left(\int_0^T \Pi_s \cdot {}^t \Pi_s ds \right) \theta \right] = 0$$

ou, de manière équivalente,

$$\theta \left(\int_0^T \Pi_s^T \cdot {}^t \Pi_s ds \right) + \left(\int_0^T \Pi_s^T \cdot {}^t \Pi_s ds \right) \theta = \Pi_T^{2\otimes} - T \cdot \Gamma .$$

C'est-à-dire, quand $x \in C$ est l'observation, on doit résoudre l'équation

$$\theta \left(\int_0^T x_s^T \cdot {}^t x_s ds \right) + \left(\int_0^T x_s^T \cdot {}^t x_s ds \right) \theta = x_T^T \cdot x_T - T \cdot \Gamma$$

ce qui est possible presque sûrement.

On aborde maintenant l'étude des propriétés asymptotiques de la famille $\{\hat{\theta}_{\Gamma, T}; T > 0\}$ d'estimateurs. On s'appuie sur les considérations suivantes :

a) Le processus $(C, S, \mu_{\theta, \Gamma}, (\eta_t^\theta)_{t \geq 0})$ défini par :

$$\eta_t^\theta = G^{-1} \Pi_t - \int_0^t G^{-1} \theta \Pi_s ds ; t \geq 0$$

est un processus de Wiener standard et, relativement à la même base de processus, on a :

$$d\Pi_t = \theta \Pi_t + G d\eta_t^\theta ; t \geq 0 , \Pi_0 \equiv 0 .$$

b) Un théorème de Lyapunov (voir par exemple F.R. GANTMACHER [22] p. 185) assure que si la matrice réelle θ est stable (i.e. ses valeurs propres ont des parties réelles strictement négatives) alors l'équation matricielle :

$$(10) \quad \theta Q + Q \cdot {}^t \theta = -\Gamma = -G \cdot {}^t G$$

admet pour solution Q_θ la matrice d'une forme quadratique définie positive. On munit alors $(\mathbb{R}^p \times C, \mathcal{B}_{\mathbb{R}^p} \otimes S)$, où $\mathcal{B}_{\mathbb{R}^p}$ désigne la tribu de Borel sur \mathbb{R}^p , de la mesure de probabilité $\tilde{\mu}_{\theta, \Gamma}$ définie par :

$$\tilde{\mu}_{\theta, \Gamma} = N(0, Q_\theta) \otimes \mu_{\theta, \Gamma}$$

où $N(0, Q_\theta)$ est la distribution gaussienne centrée de matrice de covariances Q_θ sur \mathbb{R}^p . On définit les fonctions aléatoires $(\tilde{\Pi}_t; t \geq 0)$ et $(\tilde{\eta}_t^\theta; t \geq 0)$ sur $(\mathbb{R}^p \times C, \mathcal{B}_{\mathbb{R}^p} \otimes S)$ par :

$$\begin{cases} \tilde{\Pi}_t(a, x) = \Pi_t(x) \\ \tilde{\eta}_t^\theta(a, x) = \eta_t^\theta(x) \end{cases} ; (a, x) \in \mathbb{R}^p \times C .$$

Clairement, le processus $(\tilde{\Pi}_t ; t \geq 0 ; \tilde{\mu}_{\theta, \Gamma})$ vérifie l'équation :

$$\begin{cases} d\tilde{\Pi}_t = \theta \tilde{\Pi}_t dt + G d\tilde{\eta}_t^\theta ; t \geq 0 \\ \tilde{\Pi}_0 \equiv 0 \end{cases}$$

où le processus $(\tilde{\eta}_t^\theta ; t \geq 0 ; \tilde{\mu}_{\theta, \Gamma})$ est un processus de Wiener standard.

On considère alors, relativement à la base de processus $(\mathbb{R}^p \times C, \mathcal{B}_{\mathbb{R}^p \otimes S}, \tilde{\mu}_{\theta, \Gamma})$ l'équation différentielle stochastique :

$$\begin{cases} dU_t = \theta U_t dt + G d\tilde{\eta}_t^\theta ; t \geq 0 \\ U_0 = \Pi_0^* \end{cases}$$

où Π_0^* désigne la projection de $\mathbb{R}^p \times C$ sur \mathbb{R}^p . La solution (forte) de cette équation est le processus gaussien centré défini par :

$$\Pi_t^* = e^{t\theta} (\Pi_0^* + \int_0^t e^{-s\theta} G d\tilde{\eta}_s^\theta) ; t \geq 0 .$$

Il est facile de vérifier que ce processus est stationnaire de fonction de covariance R_θ donnée par :

$$(11) \quad R_\theta(h) = e^{h\theta} Q_\theta ; h \geq 0 .$$

On démontre alors quelques lemmes préliminaires :

1.3.3. LEMME : "Soit θ une matrice stable $p \times p$ et G une matrice non singulière $p \times p$. Soit $(N, \eta, \mu, (\gamma_t ; t \geq 0))$ un processus de Wiener standard p -dimensionnel. Si $(Z_t)_{t \geq 0}$ est solution de l'équation différentielle stochastique

$$(12) \quad dU_t = \theta U_t dt + G d\gamma_t ; t \geq 0$$

de condition initiale $Z_0 \equiv 0$, et $(Z_t^*)_{t \geq 0}$ est un processus centré stationnaire au second ordre vérifiant (12), alors :

$$(i) \quad \left\{ \int_0^T Z_s^* \cdot {}^t Z_s^* ds - \int_0^T Z_s \cdot {}^t Z_s ds ; T > 0 \right\}$$

$$(ii) \left\{ \int_0^T dZ_s^* \cdot {}^t Z_s^* - \int_0^T dZ_s \cdot {}^t Z_s ; T > 0 \right\}$$

sont deux familles bornées en moyenne quadratique."

Pour la simplicité de l'exposé nous supposons dans les démonstrations que les valeurs propres de la matrice stable θ concernée sont toutes différentes ; dans le cas contraire, utilisant la décomposition canonique de Jordan de θ on pourrait conduire des raisonnements analogues.

Démonstration : Pour $t \geq 0$, Z_t peut s'écrire sous la forme :

$$Z_t = \int_0^t e^{(t-u)\theta} G d\gamma_u = Z_t^* - e^{t\theta} Z_0^*.$$

On évalue alors :

$$(13) \quad \int_0^T Z_s^* \cdot {}^t Z_s^* ds - \int_0^T Z_s \cdot {}^t Z_s ds \\ = \int_0^T e^{s\theta} Z_0^* \cdot {}^t Z_s ds + \int_0^T Z_s \cdot Z_0^* e^{s^t\theta} ds + \int_0^T e^{s\theta} Z_0^* \cdot {}^t Z_0^* \cdot e^{s^t\theta} ds.$$

On a :

$$E \left\| \int_0^T e^{s\theta} Z_0^* \cdot {}^t Z_s ds \right\|^2 \\ = \text{Tr} E \left\{ \left[\int_0^T e^{u\theta} Z_0^* \cdot {}^t Z_u du \right] \cdot \left[\int_0^T e^{s\theta} Z_0^* \cdot {}^t Z_s ds \right] \right\} \\ = \text{Tr} E \left\{ \int_0^T \int_0^T e^{u\theta} Z_0^* \cdot {}^t Z_u \cdot Z_s \cdot {}^t Z_0^* \cdot e^{s^t\theta} du ds \right\} \\ = \text{Tr} \int_0^T \int_0^T e^{u\theta} E(Z_0^* \cdot {}^t Z_u \cdot Z_s \cdot {}^t Z_0^*) e^{s^t\theta} du ds.$$

Compte-tenu de l'indépendance des éléments aléatoires Z_0^* et ${}^t Z_u \cdot Z_s$, et de ce que tout processus stationnaire au second ordre vérifiant (12) a pour fonction de covariance la fonction R_θ donnée par (11) (voir par exemple L. ARNOLD [4], Théorème 8.2.12) il vient :

$$E \left\| \int_0^T e^{s\theta} Z_0^* \cdot {}^t Z_s ds \right\|^2 = \text{Tr} \int_0^T \int_0^T e^{u\theta} Q_\theta e^{s^t\theta} E({}^t Z_u Z_s) du ds.$$

Or, utilisant le fait que $Z_t = Z_t^* - e^{t\theta} Z_0^*$, où $(Z_t^*)_{t \geq 0}$ est un processus centré stationnaire au second ordre de fonction de covariance R_θ , on obtient aisément que $E(Z_s \cdot {}^t Z_s) = Q_\theta - e^{s\theta} Q_\theta e^{s^t\theta}$.

Alors $E({}^tZ_s . Z_s) = \text{Tr} E(Z_s . {}^tZ_s) \leq \text{Tr} Q_\theta$
 puisque la matrice $e^{s\theta} . Q_\theta . e^{st\theta}$ est symétrique définie positive. Par suite, on a :

$$|E({}^tZ_u . Z_s)| \leq [E({}^tZ_u Z_u)]^{1/2} . [E({}^tZ_s . Z_s)]^{1/2} \leq \text{Tr} Q_\theta .$$

En définitive on a la majoration :

$$E \left\| \int_0^T e^{s\theta} Z_0^* {}^tZ_s ds \right\|^2 \leq (\text{Tr} Q_\theta) \int_0^T \int_0^T |\text{Tr} \{e^{u\theta} Q_\theta e^{st\theta}\}| du ds .$$

Ecrivant $\theta = C \Lambda C^{-1}$ où Λ est la matrice diagonale dont les coefficients diagonaux λ_i sont les valeurs propres de θ et C est une matrice non singulière, il vient :

$$e^{u\theta} = C e^{u\Lambda} C^{-1} ,$$

alors si $\sup\{\text{Re}(\lambda_i) ; i=1, p\} \leq \lambda < 0$, il existe une constante K telle que

$$\text{Max}_{i,j} |(e^{u\theta})_{i,j}| \leq K e^{\lambda u} ; u \geq 0 ,$$

et il existe une constante K' telle que :

$$|\text{Tr} \{e^{u\theta} Q_\theta e^{st\theta}\}| \leq K' e^{\lambda(u+s)} ; u \geq 0 , s \geq 0 .$$

On en déduit immédiatement que l'expression $E \left\| \int_0^T e^{s\theta} Z_0^* {}^tZ_s ds \right\|^2$ est majorée uniformément par rapport à T . Les deux premiers termes du membre de droite de l'égalité (13) sont donc majorés en moyenne quadratique uniformément par rapport à T .

Le troisième terme du membre de droite de l'égalité (13) est une matrice symétrique semi-définie positive. Son espérance est :

$$E \int_0^T e^{s\theta} . Z_0^* . {}^tZ_0^* e^{st\theta} = \int_0^T e^{s\theta} . Q_\theta . e^{st\theta} ds$$

dont la trace est encore majorée par une constante indépendante de T . Ainsi ce troisième terme est majoré en moyenne d'ordre 1 uniformément par rapport à T . En fait, procédant pour ce terme comme il a été fait pour le premier, on peut voir que la propriété de majoration est aussi vraie en moyenne quadratique. Ainsi l'assertion du lemme est démontrée pour (i).

On évalue maintenant :

$$(14) \quad \int_0^T dZ_s^* \cdot {}^t Z_s^* - \int_0^T dZ_s \cdot {}^t Z_s$$

$$= \theta \left\{ \int_0^T Z_s^* \cdot {}^t Z_s^* ds - \int_0^T Z_s \cdot {}^t Z_s ds \right\} + \int_0^T G d\gamma_s \cdot {}^t (Z_s^* - Z_s) .$$

D'après la première partie de la démonstration, le premier terme du second membre de l'égalité (14) est majoré uniformément par rapport à T dans $[L^2(N, \eta, \mu)]^{p^2}$. Le deuxième terme s'écrit :

$$G \int_0^T d\gamma_s \cdot {}^t Z_0^* e^{s\theta}$$

et on a :

$$E \left\| \int_0^T d\gamma_s \cdot {}^t Z_0^* e^{s\theta} \right\|^2 = p \int_0^T E \| e^{s\theta} Z_0^* \|^2 ds .$$

Cette expression est elle aussi majorée uniformément par rapport à T ce qui permet de conclure à l'assertion du lemme pour (ii). ■

1.3.4. LEMME : "Sous les hypothèses du lemme 1.3.3., les familles

$$(i) \quad \{ T^{-1/2} [\int_0^T dZ_s^* \cdot {}^t Z_s^* - \theta \int_0^T Z_s^* \cdot {}^t Z_s^* ds] ; T > 0 \}$$

et

$$(ii) \quad \{ T^{-1/2} [\int_0^T dZ_s \cdot {}^t Z_s - \theta \int_0^T Z_s \cdot {}^t Z_s ds] ; T > 0 \}$$

sont bornées en moyenne quadratique".

Démonstration : On évalue :

$$\int_0^T dZ_s^* \cdot {}^t Z_s^* - \int_0^T \theta Z_s^* \cdot {}^t Z_s^* ds = \int_0^T G d\gamma_s \cdot {}^t Z_s^* .$$

On a :

$$E \left\| \int_0^T d\gamma_s \cdot {}^t Z_s^* \right\|^2 = p \int_0^T E \| Z_s^* \|^2 ds$$

et, comme $(Z_s^* ; s \geq 0)$ est un processus centré stationnaire au second ordre,

$$\frac{1}{T^{1/2}} \left\{ \int_0^T dZ_s^* \cdot {}^t Z_s^* - \theta \int_0^T Z_s^* \cdot {}^t Z_s^* ds \right\}$$

est majoré, dans $[L^2(N, \eta, \mu)]^{p^2}$, uniformément par rapport à T . Cela

assure l'assertion du lemme pour (i).

L'assertion pour (ii) se déduit immédiatement de celle pour (i) et du lemme 1.3.3. ■

Nous sommes alors en mesure de démontrer une proposition qui assure la convergence de la famille d'estimateurs $\{\hat{\theta}_{\Gamma, T}; T > 0\}$:

1.3.5. PROPOSITION : "Si Θ est le sous-ensemble de $\mathfrak{L}(\mathbb{R}^p; \mathbb{R}^p)$ constitué des matrices stables, la matrice θ est asymptotiquement identifiable en probabilité par la famille $\{\hat{\theta}_{\Gamma, T}; T > 0\}$ de statistiques définies par (9), i.e. pour tout $\theta \in \Theta$, on a,

$$(15) \quad \mu_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} \hat{\theta}_{\Gamma, T} = \theta . "$$

Démonstration : On considère les processus $(\tilde{\Pi}_t; t \geq 0)$ et $(\Pi_t^*; t \geq 0)$ basés sur l'espace probabilisé $(\mathbb{R}^p \times C, \mathcal{B}_{\mathbb{R}^p} \otimes S, \tilde{\mu}_{\theta, \Gamma})$ définis précédemment et on pose :

$$\tilde{\theta}_{\Gamma, T}(a, x) = \hat{\theta}_{\Gamma, T}(x) \quad \text{pour tout } (a, x) \in \mathbb{R}^p \times C .$$

On a encore, compte-tenu de la définition de $\hat{\theta}_{\Gamma, T}$ par (9) :

$$(16) \quad \frac{1}{T} \left\{ \int_0^T d\tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} - \tilde{\theta}_{\Gamma, T} \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} ds \right\} = 0$$

$\tilde{\mu}_{\theta, \Gamma}$ -presque sûrement.

D'après le lemme 1.3.4. on a :

$$(17) \quad \tilde{\mu}_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \left\{ \int_0^T d\tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} - \theta \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} ds \right\} = 0 .$$

Les égalités (16) et (17) montrent que :

$$(18) \quad \tilde{\mu}_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \left\{ (\tilde{\theta}_{\Gamma, T} - \theta) \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} ds \right\} = 0 .$$

Le processus $(\Pi_t^*; t \geq 0; \tilde{\mu}_{\theta, \Gamma})$ étant gaussien centré stationnaire de fonction de covariance R_θ donnée par (11), admet une densité spectrale définie par :

$$f(\lambda) = \frac{1}{2\pi} (\theta - i\lambda \mathbf{1}_p)^{-1} \Gamma (\theta + i\lambda \mathbf{1}_p)^{-1} .$$

Un théorème ergodique (voir [45] Yu A. ROZANOV p. 157 et p. 163) assure que dans ces conditions on a :

$$\tilde{\mu}_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T \Pi_s^* \cdot t_{\Pi_s}^* ds = R_{\theta}(0) = Q_{\theta} .$$

Le lemme 1.3.3. permet donc d'affirmer que :

$$(19) \quad \tilde{\mu}_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} ds = Q_{\theta} .$$

On évalue alors la limite en probabilité de :

$$(\tilde{\theta}_{\Gamma, T}^{-\theta}) = (\tilde{\theta}_{\Gamma, T}^{-\theta}) \{T^{-1} \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} ds\} \cdot \{T^{-1} \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s} ds\}^{-1} .$$

La matrice Q_{θ} étant non singulière on obtient, compte-tenu de (18) et (19) et d'énoncés classiques concernant la convergence en probabilité (voir par exemple [12] P. BILLINGSLEY, p. 31) :

$$\tilde{\mu}_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} \tilde{\theta}_{\Gamma, T} = \theta$$

d'où découle l'égalité (15). ■

1.3.6. Remarques.

a) La proposition 1.3.5. est en fait un corollaire de la proposition I. 2.2.1. : en effet, le lemme 1.3.4. montre que

$$\mu_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \left\{ \int_0^T d\Pi_s \cdot t_{\Pi_s} - \theta \int_0^T \Pi_s \cdot t_{\Pi_s} ds \right\} = 0$$

et, lors de la dernière démonstration, nous avons montré que :

$$\mu_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \left\{ \int_0^T \Pi_s \cdot t_{\Pi_s} ds \right\} = Q_{\theta}$$

où Q_{θ} est une matrice strictement négative.

Ainsi la fonction de Log-vraisemblance $L_{\Gamma, T}$ donnée par (8) vérifie les hypothèses (E1), (E2) et (E4) de la proposition I. 2.2.1. La famille $\{\hat{\theta}_{\Gamma, T}; T > 0\}$ est donc bien l'unique famille d'estimateurs de θ qui vérifie les conclusions (i), (ii) et (iii) de cette proposition, compte-tenu de la proposition 1.3.1.

b) Dans [36] Ch.17, § 4, P.S. LIPCER et A.N. SHIRYAYEV ont considéré le cas particulier où $p = 2$ et la matrice inconnue θ est de la forme :

$$\theta = \begin{pmatrix} -\theta_1 & -\theta_2 \\ \theta_2 & -\theta_1 \end{pmatrix} ; \quad \theta_1 > 0 ; \quad -\infty < \theta_2 < +\infty .$$

Il est facile de voir qu'une telle matrice est stable ; la convergence en probabilité des estimateurs de maximum de vraisemblance de θ_1 et θ_2 résulte donc de l'étude précédente. Cet exemple est repris dans [41] A.A. NOVIKOV et [49] A.F. TARASKIN.

c) Dans le cas unidimensionnel, i.e. $p = 1$ les résultats peuvent être améliorés. En effet, on peut montrer, que l'estimateur $\hat{\theta}_{\Gamma, T}$, défini alors par :

$$\hat{\theta}_{\Gamma, T}(x) = \frac{x_T^2 - \Gamma T}{2 \int_0^T x_s^2 ds} ,$$

est en fait convergent presque sûrement, i.e. pour tout $\theta \in \mathbb{R}$:

$$\lim_{T \rightarrow +\infty} \hat{\theta}_{\Gamma, T} = \theta \quad \mu_{\theta, \Gamma} \text{-presque sûrement.}$$

Pour une démonstration, nous renvoyons à [36] Ch.17, § 3, P.S. LIPCER et A.N. SHIRYAYEV.

Nous abordons maintenant l'étude de la distribution asymptotique de la famille $\{\hat{\theta}_{\Gamma, T} ; T > 0\}$. On démontre d'abord un lemme :

1.3.7. LEMME : "Sous les hypothèses du lemme 1.3.3., on a :

$$\mathfrak{L}_{\mu} - \lim_{T \rightarrow +\infty} T^{-1/2} \int_0^T G dY_s \cdot {}^t Z_s^* = \mathfrak{L}_{\mu} - \lim_{T \rightarrow +\infty} T^{-1/2} \left\{ \int_0^T dZ_s^* \cdot Z_s^* - \theta \int_0^T Z_s^* \cdot {}^t Z_s^* ds \right\} = M^{\theta}$$

où M^{θ} est une matrice aléatoire de la loi gaussienne $N(0, \Gamma^{\theta})$ où :

$$E(M_{ij}^{\theta} M_{k\ell}^{\theta}) = \Gamma_{ijk\ell}^{\theta} = (G \cdot {}^t G)_{ik} \cdot (Q_{\theta})_{j\ell}$$

Q_q étant définie par (10)."

Ce lemme peut se déduire immédiatement de la proposition I. 3.2.2. et du lemme 1.3.3. Nous choisissons toutefois d'en donner une démonstration directe, comme nous l'avons annoncé dans le chapitre précédent.

Démonstration : Pour $q \in \mathbb{N}^*$, on pose :

$$Z_t^{(p)} = Z_{\frac{j}{q}}^* \quad \text{si} \quad \frac{j}{q} \leq t < \frac{j+1}{q}, \quad j \in \mathbb{N}$$

et

$$\varepsilon_j^{(q)} = \frac{\gamma_j}{q} - \frac{\gamma_{j-1}}{q}, \quad j \in \mathbb{N}^*.$$

On évalue, pour $(N, q) \in (\mathbb{N}^*)^2$,

$$Z(N, p) = \int_0^N G d\gamma_s \cdot t_z^{(q)} = G \sum_{j=0}^{qN-1} \varepsilon_{j+1}^{(q)} \cdot t_z^* \frac{j}{q}.$$

On peut écrire $Z_{\frac{j}{q}}^*$ sous la forme :

$$Z_{\frac{j}{q}}^* = \lambda_j^{(q)} + \mu_j^{(q)}; \quad j \in \mathbb{N}$$

où $\lambda_0^{(q)} = Z_0^*$; $\lambda_j^{(q)} = \sum_{s=0}^{j-1} e^{\frac{s}{q}\theta} G e_{j-s}^{(q)} + e^{\frac{j}{q}\theta} Z_0^*$; $j \in \mathbb{N}^*$

et $\mu_0^{(q)} = 0$; $\mu_j^{(q)} = \sum_{k=0}^{j-1} \int_{\frac{k}{q}}^{\frac{k+1}{q}} [e^{\frac{j-t}{q}\theta} - e^{\frac{j-k-1}{q}\theta}] G d\gamma_t$; $j \in \mathbb{N}^*$.

Alors, pour tout $(N, q) \in (\mathbb{N}^*)^2$ on a :

$$Z(N, q) = G \sum_{j=1}^{qN} \varepsilon_j^{(q)} t_z^{(q)} \cdot \lambda_{j-1}^{(q)} + G \sum_{j=1}^{qN} \varepsilon_j^{(q)} t_z^{(q)} \cdot \mu_{j-1}^{(q)}.$$

Il est facile de montrer que, pour tout $q \in \mathbb{N}^*$:

$$(20) \quad \mu\text{-}\lim_{N \rightarrow +\infty} N^{-1/2} \sum_{j=1}^N \varepsilon_j^{(q)} \cdot t_z^{(q)} \cdot \mu_{j-1}^{(q)} = 0.$$

Pour ce faire on vérifie que $E \left\| \sum_{j=1}^N \epsilon_j^{(q)} \cdot t_{j-1}^{(q)} \right\|^2$ est majoré par une constante indépendante de N en utilisant le fait que θ est une matrice stable.

On montre alors que, pour tout $q \in \mathbb{N}^*$, la suite $\{N^{-1/2} \sum_{j=1}^N G \epsilon_j^{(q)} \cdot t_{j-1}^{(q)} ; N \in \mathbb{N}^*\}$ converge en loi : on définit $Q^\theta(q)$ comme étant la matrice symétrique définie positive solution de l'équation matricielle :

$$(21) \quad Q - e^{\frac{\theta}{q}} Q e^{\frac{t\theta}{q}} = \frac{1}{q} G \cdot {}^t G$$

dont les hypothèses faites sur θ assurent l'existence et l'unicité (voir [3] Ch.5, T.W. ANDERSON). On considère $\{\epsilon_{-u}^{(q)} ; u \in \mathbb{N}\}$ une suite de variables aléatoires définies sur (N, \mathcal{N}, μ) (éventuellement en "augmentant" l'espace) indépendantes entre elles et indépendantes de $\{\epsilon_j^{(q)} ; j \in \mathbb{N}^*\}$ et de même loi que celles-ci. On pose alors :

$$\tilde{\lambda}_t^{(q)} = \sum_{s=0}^{\infty} e^{\frac{s\theta}{q}} G \cdot \epsilon_{t-s}^{(q)} ; t \in \mathbb{Z} .$$

Ce processus est stationnaire et tel que :

$$E(\tilde{\lambda}_t^{(q)} \cdot {}^t \tilde{\lambda}_t^{(q)}) = \sum_{s=0}^{\infty} e^{\frac{s\theta}{q}} \times \frac{G \cdot {}^t G}{q} \times e^{\frac{s \cdot t \theta}{q}} = Q^\theta(q) .$$

D'après le théorème 5.5.5. de [3] Ch.5, T.W. ANDERSON, on a :

$$\mathcal{L}_\mu - \lim_{N \rightarrow +\infty} \frac{1}{\sqrt{N}} \sum_{j=1}^N G \epsilon_j^{(q)} \cdot t_{j-1}^{(q)} = G_q^\theta$$

où G_q^θ est une matrice suivant une loi normale centrée de covariances données par :

$$E(G_q^\theta(i, j) \cdot G_q^\theta(k, \ell)) = \frac{1}{q} (G \cdot {}^t G)_{i, k} (Q^\theta(q))_{j, \ell} .$$

D'autre part (voir [3] Ch.5, § 5.5.2., T.W. ANDERSON) on a :

$$\mu - \lim_{N \rightarrow +\infty} \frac{1}{\sqrt{N}} \sum_{j=1}^N G \epsilon_j^{(q)} \cdot t(\tilde{\lambda}_t^{(q)} - \lambda_t^{(q)}) = 0 .$$

Par conséquent on a aussi :

$$(22) \quad \mathfrak{L}_\mu - \lim_{N \rightarrow +\infty} \frac{1}{\sqrt{N}} \sum_{j=1}^N G \epsilon_j^{(q)} \cdot {}^t \lambda_{j-1}^{(q)} = G_q^\theta .$$

Compte-tenu de (20) et (22) on a, pour tout $q \in \mathbb{N}^*$:

$$\mathfrak{L}_\mu - \lim_{N \rightarrow +\infty} N^{-1/2} \sum_{j=1}^N G \epsilon_j^{(q)} \cdot {}^t Z_{\frac{(j-1)}{q}}^* = G_q^\theta$$

puis,

$$(23) \quad \mathfrak{L}_\mu - \lim_{N \rightarrow +\infty} N^{-1/2} Z(N, q) = q^{1/2} G_q^\theta = M_q^\theta$$

où M_q^θ est donc une matrice suivant une loi normale centrée de covariances données par :

$$E(M_q^\theta(i, j) \cdot M_q^\theta(k, \ell)) = (G \cdot {}^t G)_{i, k} (Q^\theta(q))_{j, \ell} .$$

On écrit maintenant :

$$\begin{aligned} U_N &= N^{-1/2} \int_0^N G d\gamma_s \cdot {}^t Z_s^* = N^{-1/2} \int_0^N G d\gamma_s \cdot {}^t Z_s^{(q)} + N^{-1/2} \int_0^N G d\gamma_s \cdot {}^t (Z_s^* - Z_s^{(q)}) \\ &= N^{-1/2} Z(N, q) + N^{-1/2} V(N, q) . \end{aligned}$$

D'après un lemme de BERNSTEIN (voir par exemple [25] p. 242, E.J. HANNAN) on aura :

$$(24) \quad \mathfrak{L}_\mu - \lim_{N \rightarrow +\infty} N^{-1/2} \int_0^N G d\gamma_s \cdot {}^t Z_s^* = M^\theta ,$$

où M^θ est précisée dans l'énoncé du lemme ; si on montre d'une part que la suite $\{ \|\sqrt{\frac{q}{N}} V(N, q)\|^2 ; q \in \mathbb{N}^* \}$ est bornée uniformément par rapport à N et d'autre part que : $\lim_{q \rightarrow +\infty} Q^\theta(q) = Q_\theta$.

Le premier point s'acquiert aisément en vertu de la stationnarité du processus (Z_t^*) . Pour assurer le second point, remarquons que l'équation (21) peut se réécrire sous la forme :

$$(21)' \quad q(e^{-\frac{\theta}{q}} - \mathbb{1}_p) Q - Q q(e^{-\frac{\theta}{q}} - \mathbb{1}_p) = e^{-\frac{\theta}{q}} G \cdot {}^t G .$$

Les coefficients $q(e^{-\frac{\theta}{q}} - \mathbb{1}_p)Q$, $q(e^{-\frac{t_\theta}{q}} - \mathbb{1}_p)$ et $e^{-\frac{\theta}{q}}$ de l'équation (21)' tendent respectivement, lorsque $q \rightarrow +\infty$, vers $-\theta$, t_θ et $\mathbb{1}_p$. Or l'équation :

$$-\theta Q - Q \cdot t_\theta = G \cdot t_G$$

n'est autre que l'équation (10) définissant la matrice Q_θ . Compte-tenu de l'unicité des solutions de (10) et (21) on a bien :

$$\lim_{q \rightarrow +\infty} Q^\theta(q) = Q_\theta.$$

En définitive, nous avons démontré l'égalité (24). On évalue alors :

$$\begin{aligned} T^{-1/2} \int_0^T d\gamma_s \cdot t_{Z_s}^* &= [T]^{-1/2} \int_0^{[T]} d\gamma_s \cdot t_{Z_s}^* + (T^{-1/2} - [T]^{-1/2}) \int_0^{[T]} d\gamma_s \cdot t_{Z_s}^* \\ &\quad + T^{-1/2} \int_{[T]}^T d\gamma_s \cdot t_{Z_s}^* \end{aligned}$$

où $[T]$ désigne la partie entière de T .

On a, d'une part :

$$\|(T^{-1/2} - [T]^{-1/2}) \left(\int_0^{[T]} d\gamma_s \cdot t_{Z_s}^* \right)\| \leq \frac{C}{[T]} \left\| \int_0^{[T]} d\gamma_s \cdot t_{Z_s}^* \right\|$$

et d'autre part,

$$E \left\| \int_{[T]}^T d\gamma_s \cdot t_{Z_s}^* \right\|^2 = p \int_{[T]}^T E \|Z_s^*\|^2 ds$$

qui est majoré uniformément par rapport à T .

Alors, compte-tenu de (i) lemme 1.3.4. et de (24), on a encore :

$$\mathfrak{L} - \lim_{T \rightarrow +\infty} T^{-1/2} \int_0^T G d\gamma_s \cdot t_{Z_s}^* = M^\theta$$

ce qui achève la démonstration du lemme. ■

Nous pouvons maintenant démontrer la proposition :

1.3.8. PROPOSITION : "Sous les hypothèses de la proposition 1.3.5., la famille d'estimateurs $\{\hat{\theta}_{\Gamma, T}^* ; T > 0\}$ définie par (9) est asymptotiquement normalement distribuée ; de manière précise on a, pour tout

$\theta \in \Theta$:

$$\mathfrak{L}_{\mu, \Gamma}^{\theta} - \lim_{T \rightarrow +\infty} T^{1/2} (\hat{\theta}_{\Gamma, T}^{-\theta}) = K^{\theta}$$

où K^{θ} est une matrice gaussienne centrée de covariances données par :

$$E(K_{ij}^{\theta} \cdot K_{k\ell}^{\theta}) = \Gamma_{ik} \cdot (Q_{\theta}^{-1})_{j\ell}$$

où Q_{θ} est définie par (10)."

Démonstration : On considère à nouveau les processus $(\tilde{\Pi}_t; t \geq 0)$, $(\tilde{\Pi}_t^*; t \geq 0)$ et la famille $\{\tilde{\theta}_{\Gamma, T}; T > 0\}$ utilisés dans la démonstration de la proposition 1.3.5. Les lemmes 1.3.3. et 1.3.7. permettent de montrer aisément que :

$$\mathfrak{L}_{\tilde{\mu}, \Gamma}^{\theta} - \lim_{T \rightarrow +\infty} T^{-1/2} \left\{ \int_0^T d\tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s}^{\sim} - \theta \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s}^{\sim} ds \right\} = M^{\theta}.$$

Compte-tenu de la définition de $\tilde{\theta}_{\Gamma, T}$ on a aussi :

$$\mathfrak{L}_{\tilde{\mu}, \Gamma}^{\theta} - \lim_{T \rightarrow +\infty} T^{-1/2} \left\{ (\tilde{\theta}_{\Gamma, T}^{-\theta}) \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s}^{\sim} ds \right\} = M^{\theta}$$

où, d'après la démonstration de la proposition 1.3.5.

$$\tilde{\mu}_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T \tilde{\Pi}_s \cdot t_{\tilde{\Pi}_s}^{\sim} ds = Q_{\theta}.$$

Il s'en suit que :

$$\mathfrak{L}_{\tilde{\mu}, \Gamma}^{\theta} - \lim_{T \rightarrow +\infty} T^{1/2} (\tilde{\theta}_{\Gamma, T}^{-\theta}) Q_{\theta} = M^{\theta}$$

puis :

$$\mathfrak{L}_{\mu, \Gamma}^{\theta} - \lim_{T \rightarrow +\infty} T^{1/2} (\hat{\theta}_{\Gamma, T}^{-\theta}) Q_{\theta} = M^{\theta}.$$

Enfin, la matrice Q_{θ} étant inversible, on obtient immédiatement que :

$$\mathfrak{L}_{\mu, \Gamma}^{\theta} - \lim_{T \rightarrow +\infty} T^{1/2} (\hat{\theta}_{\Gamma, T}^{-\theta}) = K^{\theta} \quad \blacksquare$$

1.3.9. Remarques.

a) Les variances et covariances de la loi limite peuvent être estimées de façon convergente (en probabilité) puisque :

$$\mu_{\theta, \Gamma} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T \Pi_s \cdot {}^t \Pi_s = Q_{\theta} .$$

b) Les résultats obtenus sont de même nature que les résultats classiques concernant le modèle autorégressif des séries chronologiques tels qu'on les trouve, par exemple, exposés dans [3] Ch.5, T.W. ANDERSON

§ 2. - ESTIMATION DANS LE MODELE LINEAIRE. CAS DE L'OBSERVATION EN TEMPS DISCRET.

Nous supposons ici que la matrice $G \cdot {}^t G$ est complètement connue, éventuellement parce qu'elle aurait été calculée, selon la proposition 1.2.1., à l'aide d'une observation en temps continu sur un court intervalle de temps. Sans perdre de généralité nous supposerons même que $G \cdot {}^t G = \mathbb{1}_p$.

D'après la proposition 1.3.1., l'estimateur de maximum de vraisemblance de θ au vu de l'observation sur $[0, T]$ est alors donnée par :

$$\hat{\theta}_T = \left[\int_0^T d\Pi_s \cdot {}^t \Pi_s \right] \left[\int_0^T \Pi_s \cdot {}^t \Pi_s ds \right]^{-1}$$

où l'intégrale stochastique est définie relativement à μ_{*}^T . En pratique, même si l'observation a lieu en temps continu, dans le calcul de $\hat{\theta}_T$, l'intégrale stochastique $\int_0^T d\Pi_s \cdot {}^t \Pi_s$ devra être approchée par une somme finie appropriée, de la forme :

$$\sum_{i=1}^N (\Pi_{t_i} - \Pi_{t_{i-1}}) \cdot {}^t \Pi_{t_{i-1}} ,$$

de même que l'intégrale ordinaire devra être approchée par une somme de Riemann du type :

$$\sum_{i=1}^N (t_i - t_{i-1}) \Pi_{t_{i-1}} \cdot {}^t \Pi_{t_{i-1}}$$

où $\{t_0, t_1, \dots, t_N\}$ définit une subdivision de $[0, T]$. Une question naturelle se pose alors : l'estimateur résultant de ces approximations,

$$\hat{\theta}_{N,T} = \left[\sum_{i=1}^N (\Pi_{t_i} - \Pi_{t_{i-1}}) \cdot t_{\Pi_{t_{i-1}}} \right] \left[\sum_{i=1}^N (t_i - t_{i-1}) \Pi_{t_{i-1}} \cdot t_{\Pi_{t_{i-1}}} \right]^{-1}$$

est-il le "meilleur" estimateur que l'on puisse obtenir au vu de l'échantillonnage en temps discret aux instants t_0, t_1, \dots, t_N ? On est alors conduit à étudier le problème de l'estimation du paramètre de dérivée θ au vu d'une observation en temps discret d'un processus solution de l'équation :

$$(1') \quad dX_t = \theta X_t + dW_t \quad ; \quad t \geq 0, \quad X_0 = 0.$$

Les résultats de ce paragraphe ont été présentés dans [33]

A. LE BRETON.

2.1. La fonction de Log-vraisemblance pour l'estimation basée sur un échantillonnage discret.

La structure statistique pour l'estimation de θ au vu d'une observation aux instants $t_0^N, t_1^N, \dots, t_{K_N}^N$ où $0 = t_0^N < t_1^N < \dots < t_{K_N}^N = T$, est :

$$(C, S_T^N, \{\mu_\theta^{T,N} ; \theta \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p)\})$$

où :

- S_T^N est la sous-tribu de S_T engendrée par la famille $\{\Pi_{t_i}^N ; 0 \leq i \leq K_N\}$;
- pour tout $\theta \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p)$, $\mu_\theta^{T,N}$ est la restriction de $P_{\theta, \mathbb{1}_p}^T$ à S_T^N .

Nous évaluons d'abord la fonction de vraisemblance $\frac{d\mu_\theta^{T,N}}{d\mu_*^{T,N}}$ où $\mu_*^{T,N}$ désigne la restriction de μ_* à S_T^N .

2.1.1. LEMME : "Pour tout $\theta \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^p)$ les mesures $\mu_\theta^{T,N}$ et $\mu_*^{T,N}$ sont équivalentes, la dérivée de Radon-Nikodym étant donnée par :

$$\frac{d\mu_{\theta}^{T,N}}{d\mu_*^{T,N}} = \left[\prod_{i=1}^{K_N} \frac{\det \Sigma_{\theta,i}^N}{(t_i^N - t_{i-1}^N)^p} \right]^{-1/2} \cdot \exp \left\{ -\frac{1}{2} \sum_{i=1}^{K_N} t_i^N \left(\begin{matrix} (t_i^N - t_{i-1}^N) \theta \\ \Pi_{t_i^N} - e \\ \Pi_{t_{i-1}^N} \end{matrix} \right) \right.$$

$$\left. \left(\Sigma_{\theta,i}^N \right)^{-1} \cdot \left(\begin{matrix} (t_i^N - t_{i-1}^N) \theta \\ \Pi_{t_i^N} - e \\ \Pi_{t_{i-1}^N} \end{matrix} \right) + \frac{1}{2} \sum_{i=1}^{K_N} \frac{\left(\begin{matrix} \Pi_{t_i^N} - \Pi_{t_{i-1}^N} \\ \Pi_{t_i^N} - \Pi_{t_{i-1}^N} \end{matrix} \right) \left(\begin{matrix} \Pi_{t_i^N} - \Pi_{t_{i-1}^N} \\ \Pi_{t_i^N} - \Pi_{t_{i-1}^N} \end{matrix} \right)}{t_i^N - t_{i-1}^N} \right\}$$

où $\Sigma_{\theta,i}^N = \int_0^{t_i^N - t_{i-1}^N} e^{u\theta} \cdot e^{u^t \theta} du$; $i = \overline{1, K_N}$."

Démonstration : L'équivalence des mesures $\mu_{\theta}^{T,N}$ et $\mu_*^{T,N}$ résulte immédiatement de celle des mesures $P_{\theta, \mathbb{1}_p}^T$ et μ_*^T . Nous calculons alors la densité de la distribution de probabilité $P_{(X_{t_1^N}, \dots, X_{t_{K_N}^N})}$ du vecteur aléatoire $(X_{t_1^N}, \dots, X_{t_{K_N}^N})$ par rapport à celle de

$(W_{t_1^N}, \dots, W_{t_{K_N}^N})$, soit $P_{(W_{t_1^N}, \dots, W_{t_{K_N}^N})}$.

Pour ce faire on calcule les densités respectives de ces distributions par rapport à la mesure de Lebesgue sur $(\mathbb{R}^p)^{K_N}$, soit λ_p^N .

La solution de (1)' est un processus gaussien centré donné par :

$$X_t = \int_0^t e^{(t-s)\theta} dW_s \quad ; \quad t \geq 0$$

de covariance :

$$r(s, t) = e^{s\theta} \left[\int_0^s e^{-u\theta} \cdot e^{-u^t \theta} du \right] e^{t\theta} \quad \text{pour } 0 \leq s \leq t .$$

On peut écrire :

$$X_{t_{i+1}^N} = e^{(t_{i+1}^N - t_i^N)\theta} X_{t_i^N} + e^{t_{i+1}^N \theta} \int_{t_i^N}^{t_{i+1}^N} e^{-s\theta} dW_s \quad \text{pour } i = \overline{0, K_N - 1} .$$

Posons :

$$\epsilon_i = e^{t_i^N \theta} \int_{t_{i-1}^N}^{t_i^N} e^{-s\theta} dW_s \quad \text{pour } i = \overline{1, K_N} .$$

Alors :

$$\epsilon_1 = X_{t_1^N} \quad ; \quad \epsilon_2 = X_{t_2^N} - e^{\frac{(t_2^N - t_1^N)\theta}{2}} X_{t_1^N} \quad ; \quad \dots \quad ; \quad \epsilon_{K_N} = X_{t_{K_N}^N} - e^{\frac{(t_{K_N}^N - t_{K_N-1}^N)\theta}{2}} X_{t_{K_N-1}^N}$$

où la loi $P_{(\epsilon_1, \dots, \epsilon_{K_N})}$ du vecteur aléatoire $(\epsilon_1, \dots, \epsilon_{K_N})$ est gaussienne centrée et telle que :

$$E(\epsilon_i \cdot \epsilon_j) = \begin{cases} e^{t_i^N \theta} \left[\int_{t_{i-1}^N}^{t_i^N} e^{-u\theta} \cdot e^{-u\theta} du \right] e^{t_i^N \theta} & \text{si } i=j \\ 0 & \text{si } i \neq j \end{cases}$$

On a :

$$E(\epsilon_i \cdot \epsilon_i) = \int_0^{t_i^N - t_{i-1}^N} e^{u\theta} \cdot e^{u\theta} du = \Sigma_{\theta, i}^N$$

Il s'en suit que :

$$\frac{dP_{(\epsilon_1, \dots, \epsilon_{K_N})}}{d\lambda_p^N}(x_1, \dots, x_{K_N}) = \prod_{i=1}^{K_N} f_{\epsilon_i}(x_i)$$

$$\text{où } f_{\epsilon_i}(x_i) = (2\pi)^{-p/2} (\det \Sigma_{\theta, i}^N)^{-1/2} \exp\left\{-\frac{1}{2} x_i^T \Sigma_{\theta, i}^N^{-1} x_i\right\}$$

Par suite :

$$\frac{dP_{(X_{t_1^N}, \dots, X_{t_{K_N}^N})}}{d\lambda_p^N}(x_1, \dots, x_{K_N}) = (2\pi)^{-\frac{pK_N}{2}} \left(\prod_{i=1}^{K_N} \det \Sigma_{\theta, i}^N \right)^{-1/2} \times \exp \left\{ -\frac{1}{2} \sum_{i=1}^{K_N} t_i^N \left(x_i - e^{\frac{(t_i^N - t_{i-1}^N)\theta}{2}} x_{i-1} \right)^T \left(\Sigma_{\theta, i}^N \right)^{-1} \left(x_i - e^{\frac{(t_i^N - t_{i-1}^N)\theta}{2}} x_{i-1} \right) \right\}$$

où $x_0 \equiv 0$.

Le vecteur aléatoire $(W_{t_1^N}, W_{t_2^N} - W_{t_1^N}, \dots, W_{t_{K_N}^N} - W_{t_{K_N-1}^N})$ a des composantes indépendantes gaussiennes centrées et de covariances

$(t_i^N - t_{i-1}^N) \mathbb{1}_p$, $i = \overline{1, K_N}$. Sa densité par rapport à λ_p^N est donc :

$$(2\pi)^{-\frac{pK_N}{2}} \left(\prod_{i=1}^N (t_i^N - t_{i-1}^N) \right)^{-\frac{p}{2}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{K_N} \frac{t_{x_i} \cdot x_i}{t_i^N - t_{i-1}^N} \right\} .$$

On a alors :

$$\frac{dP_{(W_{t_1}^N, \dots, W_{t_{K_N}}^N)}(x_1, \dots, x_{K_N})}{d\lambda_p^N} = (2\pi)^{-\frac{pK_N}{2}} \left(\prod_{i=1}^N (t_i^N - t_{i-1}^N) \right)^{-\frac{p}{2}} \times \exp \left\{ -\frac{1}{2} \sum_{i=1}^{K_N} \frac{t(x_i - x_{i-1})(x_i - x_{i-1})}{t_i^N - t_{i-1}^N} \right\} .$$

Le résultat annoncé s'obtient alors en écrivant :

$$\frac{dP_{(X_{t_1}^N, \dots, X_{t_{K_N}}^N)}}{dP_{(W_{t_1}^N, \dots, W_{t_{K_N}}^N)}} = \frac{dP_{(X_{t_1}^N, \dots, X_{t_{K_N}}^N)}}{d\lambda_p^N} \times \left[\frac{dP_{(W_{t_1}^N, \dots, W_{t_{K_N}}^N)}}{d\lambda_p^N} \right]^{-1} .$$

La fonction de Log-vraisemblance pour l'estimation de θ est donc :

$$\begin{aligned} \tilde{L}_{N,T}(\theta, \cdot) &= -\frac{1}{2} \left[\sum_{i=1}^{K_N} \text{Log det } \Sigma_{\theta,i}^N - p \sum_{i=1}^{K_N} \text{Log}(t_i^N - t_{i-1}^N) \right] \\ &\quad - \frac{1}{2} \sum_{i=1}^{K_N} t_{(\Pi_{t_i}^N - e^{(t_i^N - t_{i-1}^N)\theta} \Pi_{t_{i-1}^N})} (\Sigma_{\theta,i}^N)^{-1} (e^{(t_i^N - t_{i-1}^N)\theta} \Pi_{t_i}^N - \Pi_{t_{i-1}^N}) \\ &\quad + \frac{1}{2} \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^{-1} \cdot t_{(\Pi_{t_i}^N - \Pi_{t_{i-1}^N})} \cdot (\Pi_{t_i}^N - \Pi_{t_{i-1}^N}) . \end{aligned}$$

Soit $\{\Pi_N\}_{N \geq 1}$ une suite de subdivisions $\Pi_N = \{0 = t_0^N < t_1^N < \dots < t_{K_N}^N = T\}$

telle que, posant $\delta_N = \text{Max}_{i=1, \dots, K_N} (t_i^N - t_{i-1}^N)$, on ait $\lim_{N \rightarrow +\infty} \delta_N = 0$. Nous

envisageons une approximation asymptotique ($N \rightarrow +\infty$) de $\tilde{L}_{N,T}(\theta, \cdot)$.

Définissons :

$$\hat{L}_{N,T}(\theta, \cdot) = \sum_{i=1}^{K_N} t_{\Pi_{t_{i-1}}^N} \cdot t_{\theta \cdot (\Pi_{t_i}^N - \Pi_{t_{i-1}}^N)} - \frac{1}{2} \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \cdot t_{\Pi_{t_{i-1}}^N} \cdot t_{\theta \cdot \Pi_{t_{i-1}}^N}$$

qui, en fait, apparaît comme une approximation naturelle de

$$L_T(\theta, \cdot) = L_{T, \mathbb{P}}(\theta, \cdot) \text{ , i.e. } L_T(\theta, \cdot) = \int_0^T \langle \theta \Pi_t, d\Pi_t \rangle - \frac{1}{2} \int_0^T \langle \theta \Pi_t, \theta \Pi_t \rangle dt .$$

Nous allons étudier les propriétés asymptotiques des différences :

$$\tilde{L}_{N,T}(\theta, \cdot) - \hat{L}_{N,T}(\theta, \cdot) \text{ et } \hat{L}_{N,T}(\theta, \cdot) - L_T(\theta, \cdot)$$

quand N tend vers l'infini.

$\hat{L}_{N,T}^{(i)}$ (resp. $\tilde{L}_{N,T}^{(i)}$, $L_T^{(i)}$) désignera, pour $i = 0, 1, 2$, la dérivée d'ordre i , par rapport à θ , de $\tilde{L}_{N,T}$ (resp. $\hat{L}_{N,T}$, L_T).

2.2. La fonction de Log-vraisemblance approchée pour l'estimation basée sur un échantillonnage discret.

Dans la proposition suivante nous décrivons les propriétés de l'approximation de $\tilde{L}_{N,T}$ par $\hat{L}_{N,T}$:

2.2.1. PROPOSITION : "Pour toute constante $K > 0$ et pour $i = 0, 1, 2$, il existe une suite $\{\varepsilon_N^i(K) ; N \geq 1\}$ de variables aléatoires positives telles que, pour tout $\theta^* \in \mathcal{L}(\mathbb{R}^p ; \mathbb{R}^p)$:

(i) $\{\varepsilon_N^i(K) ; N \geq 1\}$ est bornée en probabilité $\mu_{\theta^*}^T$, i.e.

$$\lim_{A \rightarrow +\infty} \limsup_{N \rightarrow +\infty} P[|\varepsilon_N^i(K)| > A] = 0$$

(ii) $\sup_{\{\theta \in \mathcal{L}(\mathbb{R}^p ; \mathbb{R}^p) : |\theta| \leq K\}} |\tilde{L}_{N,T}^{(i)}(\theta, \cdot) - \hat{L}_{N,T}^{(i)}(\theta, \cdot)| \leq \delta_N^{1/2} \varepsilon_N^i(K)$,
 $\mu_{\theta^*}^T$ -presque sûrement

où pour $A \in \mathcal{L}(\mathbb{R}^k ; \mathbb{R}^l)$, $|A| = \text{Max}|A_{ij}|$."

La démonstration est développée au cours d'une succession de lemmes.

2.2.1. LEMME : "Pour $h > 0$, on définit $\Sigma_{\theta, h}$ par :

$$\Sigma_{\theta, h} = \int_0^h e^{u\theta} \cdot e^{u^t \theta} du .$$

Alors, pour h suffisamment petit, on peut écrire :

- (i) $\det \Sigma_{\theta, h} = h^n [1 + h \operatorname{tr} \theta + h^2 \gamma_1(h, \theta)]$
- (ii) $\operatorname{Log} \det \Sigma_{\theta, h} = n \operatorname{Log} h + h \operatorname{tr} \theta + h^2 \gamma_2(h, \theta)$
- (iii) $\Sigma_{\theta, h}^{-1} = h^{-1} [\mathbb{1}_p - \frac{h}{2}(\theta + {}^t\theta) + h^2 \Delta(h, \theta)]$

où $\sup_{|\theta| \leq K} |\gamma_i(h, \theta)| \leq \gamma_i(h, K)$, $i = \overline{1, 2}$

et $\sup_{|\theta| \leq K} |\Delta(h, \theta)| \leq \gamma_3(h, K)$

avec $\lim_{h \rightarrow 0} \gamma_i(h, K) = \gamma_i(K)$, $i = \overline{1, 3}$."

Démonstration : $\Sigma_{\theta, h} = h \mathbb{1}_p + \frac{h^2}{2}(\theta + {}^t\theta) + \frac{h^3}{3!} \left(\frac{d^2}{du^2} [e^{u\theta} \cdot e^{u{}^t\theta}] \right) u = h_{\theta}$
avec $0 \leq h_{\theta} \leq h$.

Alors $\Sigma_{\theta, h} = h \left[\mathbb{1}_p + \frac{h}{2}(\theta + {}^t\theta) + h^2 \Delta_0(h, \theta) \right]$

où $\sup_{|\theta| \leq K} |\Delta_0(h, \theta)| \leq \gamma_0(h, K)$ avec $\lim_{h \rightarrow 0} \gamma_0(h, K) = \gamma_0(K)$.

On a aussi :

$$\det \Sigma_{\theta, h} = h^n \prod_{i=1}^p (1 + h \lambda_{\theta, h}^i)$$

où $\{\lambda_{\theta, h}^i ; i = \overline{1, p}\}$ est l'ensemble des valeurs propres de $\frac{1}{2}(\theta + {}^t\theta + 2h\Delta_0(h, \theta)) = \Gamma(h, \theta)$.

Alors, développant le produit, on obtient :

$$\det \Sigma_{\theta, h} = h^n [1 + h \operatorname{tr} \theta + h^2 \gamma_1(h, \theta)]$$

où $\gamma_1(h, \theta)$ vérifie la propriété requise en (i).

Pour h suffisamment petit, on déduit aisément l'assertion (ii).

Pour démontrer l'assertion (iii), on écrit :

$$\begin{aligned} \Sigma_{\theta, h}^{-1} &= \{h[\mathbb{1}_p + \frac{h}{2}(\theta + {}^t\theta) + h^2 \Delta_0(h, \theta)]\}^{-1} \\ &= h^{-1} [\mathbb{1}_p + \frac{h}{2}(\theta + {}^t\theta) + h^2 \Delta_0(h, \theta)]^{-1} \\ &= h^{-1} [\mathbb{1}_p - \frac{h}{2}(\theta + {}^t\theta) + h^2 \Delta(h, \theta)] \end{aligned}$$

où $\Delta(h, \theta) = -\Delta_0(h, \theta) + \Gamma^2(h, \theta) [1_p - h\Gamma(h, \theta) + h^2\Gamma(h, \theta) + \dots]$

et $\sup_{|\theta| \leq K} |\Delta(h, \theta)| \leq \gamma_0(h, K) + n^2 [K + h\gamma_0(h, K)]^2 [1 - h(K + h\gamma_0(h, K))]^{-1}$
 $= \gamma_3(h, K) .$

2.2.3. LEMME : "On a :

$$\left| \sum_{i=1}^{K_N} (\pi_{t_i^N} - \pi_{t_{i-1}^N})^{2\otimes} - T1_p \right| \leq \delta_N^{1/2} U_N ,$$

où $\{U_N ; N \geq 1\}$ est une suite de variables aléatoires bornée en probabilité μ_θ^T pour tout $\theta \in \mathfrak{L}(\mathbb{R}^p ; \mathbb{R}^p)$."

Démonstration : Posons :

$$\sum_{i=1}^{K_N} (\pi_{t_i^N} - \pi_{t_{i-1}^N})^{2\otimes} - T1_p = \delta_N^{1/2} \tilde{U}_N ,$$

$U_N = |\tilde{U}_N|$, $t_{\pi_s} = (\pi_s^1, \dots, \pi_s^p)$ et $\delta_{k, \ell} = \begin{cases} 1 & \text{si } k = \ell \\ 0 & \text{si } k \neq \ell \end{cases}$ $k, \ell \in \{1, \dots, p\}$

Nous montrons que chaque terme de \tilde{U}_N est bornée en probabilité μ_θ^T .
 Supposons d'abord que $\theta = 0$, i.e. $\mu_\theta^T = \mu_*^T$. Alors :

$$E_{\mu_*^T} \left| \sum_{i=1}^{K_N} (\pi_{t_i^N}^k - \pi_{t_{i-1}^N}^k) (\pi_{t_i^N}^\ell - \pi_{t_{i-1}^N}^\ell) - T\delta_{k\ell} \right|^2$$

$$= \begin{cases} 2 \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^2 & \text{si } k = \ell \\ \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^2 & \text{si } k \neq \ell \end{cases}$$

$$\leq 2T\delta_N$$

et le résultat est acquis.

Maintenant, pour θ quelconque dans $\mathfrak{L}(\mathbb{R}^p ; \mathbb{R}^p)$, relativement à μ_θ^T , on peut écrire :

$$\pi_{t_i^N} - \pi_{t_{i-1}^N} = \int_{t_{i-1}^N}^{t_i^N} \theta \pi_s ds + \xi_{t_i^N}^\theta - \xi_{t_{i-1}^N}^\theta$$

où (ξ_t^θ) est un processus de Wiener. Alors, posant $t_{\xi_s^\theta} = (\xi_s^{\theta,1}, \dots, \xi_s^{\theta,p})$ et $t(\theta \Pi_s) = ((\theta \Pi_s)_1, \dots, (\theta \Pi_s)_p)$, on a :

$$\begin{aligned}
& E_{\mu_\theta^T} \left| \sum_{i=1}^{K_N} (\Pi_{t_i^N}^k - \Pi_{t_{i-1}^N}^k) (\Pi_{t_i^N}^\ell - \Pi_{t_{i-1}^N}^\ell) - T \delta_{k,\ell} \right|^2 \\
& \leq 4 E_{\mu_\theta^T} \left\{ \left| \sum_{i=1}^{K_N} (\xi_{t_i^N}^{\theta,k} - \xi_{t_{i-1}^N}^{\theta,k}) (\xi_{t_i^N}^{\theta,\ell} - \xi_{t_{i-1}^N}^{\theta,\ell}) - T \delta_{k,\ell} \right|^2 \right. \\
& \quad + \left| \sum_{i=1}^{K_N} \left(\int_{t_{i-1}^N}^{t_i^N} (\theta \Pi_s)_k ds \right) \left(\int_{t_{i-1}^N}^{t_i^N} (\theta \Pi_s)_\ell ds \right) \right|^2 \\
& \quad + \left| \sum_{i=1}^{K_N} \left(\int_{t_{i-1}^N}^{t_i^N} (\theta \Pi_s)_k ds \right) (\xi_{t_i^N}^\ell - \xi_{t_{i-1}^N}^\ell) \right|^2 \\
& \quad \left. + \left| \sum_{i=1}^{K_N} \left(\int_{t_{i-1}^N}^{t_i^N} (\theta \Pi_s)_\ell ds \right) (\xi_{t_i^N}^k - \xi_{t_{i-1}^N}^k) \right|^2 \right\} \\
& \leq 8 T \delta_N + 4 \delta_N^2 E_{\mu_\theta^T} \left[\left(\int_0^T (\theta \Pi_s)_k^2 ds \right) \left(\int_0^T (\theta \Pi_s)_\ell^2 ds \right) \right] \\
& \quad + 4 \delta_N \left[E_{\mu_\theta^T} \left(\int_0^T (\theta \Pi_s)_k^2 ds \right)^2 \right]^{1/2} \left[E_{\mu_\theta^T} \left(\sum_{i=1}^{K_N} |\xi_{t_i^N}^{\theta,\ell} - \xi_{t_{i-1}^N}^{\theta,\ell}|^2 \right)^2 \right]^{1/2} \\
& \quad + 4 \delta_N \left[E_{\mu_\theta^T} \left(\int_0^T (\theta \Pi_s)_\ell^2 ds \right)^2 \right]^{1/2} \left[E_{\mu_\theta^T} \left(\sum_{i=1}^{K_N} |\xi_{t_i^N}^{\theta,k} - \xi_{t_{i-1}^N}^{\theta,k}|^2 \right)^2 \right]^{1/2} .
\end{aligned}$$

Relativement à μ_θ^T , (Π_t) (resp. (ξ_t^θ)) est processus gaussien (resp. de Wiener) ; aussi, chaque terme de la forme $E_{\mu_\theta^T}(\cdot)$ est fini. ■

Nous sommes alors en mesure de démontrer l'assertion de la proposition 2.2.1. pour $i=0$. On peut aisément exprimer la différence $\tilde{L}_{N,T}(\rho, \cdot) - \hat{L}_{N,T}(\theta, \cdot)$ sous la forme :

$$\begin{aligned}
 (25) \quad & \tilde{L}_{N,T}(\theta, \cdot) - \hat{L}_{N,T}(\theta, \cdot) \\
 &= -\frac{1}{2} \left[\sum_{i=1}^{K_N} \text{Log det } \Sigma_{\theta,i}^N - p \sum_{i=1}^{K_N} \text{Log}(t_i^N - t_{i-1}^N) - T \text{tr } \theta \right] \\
 &+ \frac{T}{2} \text{tr} \left[\theta \left\{ \frac{1}{T} \sum_{i=1}^{K_N} (\Pi_{t_i^N} - \Pi_{t_{i-1}^N})^{2\otimes} - \mathbb{1}_p \right\} \right] \\
 &- \frac{1}{2} \text{tr} \left[\sum_{i=1}^{K_N} \left\{ (\Sigma_{\theta,i}^N)^{-1} - (t_i^N - t_{i-1}^N)^{-1} \mathbb{1}_p + \theta \right\} (\Pi_{t_i^N} - \Pi_{t_{i-1}^N})^{2\otimes} \right] \\
 &- \frac{1}{2} \text{tr} \left[\sum_{i=1}^{K_N} \left\{ (\mathbb{1}_p - e^{(t_i^N - t_{i-1}^N)\theta}) (\Sigma_{\theta,i}^N)^{-1} (\mathbb{1}_p - e^{(t_i^N - t_{i-1}^N)\theta}) - e^{\theta(t_i^N - t_{i-1}^N)} \right\} \Pi_{t_{i-1}^N}^{2\otimes} \right] \\
 &- \text{tr} \left[\sum_{i=1}^{K_N} \left\{ (\Sigma_{\theta,i}^N)^{-1} (\mathbb{1}_p - e^{(t_i^N - t_{i-1}^N)\theta}) + \theta \right\} \Pi_{t_{i-1}^N} \cdot \Pi_{t_i^N} - \Pi_{t_{i-1}^N} \right].
 \end{aligned}$$

2.2.4. LEMME : "Pour toute constante $K > 0$, il existe une suite $\{\varepsilon_N^0(K) ; N \geq 1\}$ de variables aléatoires positives telle que, pour tout $\theta^* \in \mathfrak{L}(\mathbb{R}^p ; \mathbb{R}^p)$:

- (i) $\{\varepsilon_N^0(K) ; N \geq 1\}$ est bornée en $\mu_{\theta^*}^T$ probabilité
- (ii) $\sup_{|\theta| \leq K} |\tilde{L}_{N,T}(\theta, \cdot) - \hat{L}_{N,T}(\theta, \cdot)| \leq \delta_N^{1/2} \varepsilon_N^0(K)$
 $\mu_{\theta^*}^T$ -presque sûrement."

Démonstration : On désigne par $E_{N,\theta}^i$, $i = \overline{1,5}$ les termes successifs de la somme dans (25). De $\Sigma_{\theta,i}^N = \Sigma_{\theta, t_i^N - t_{i-1}^N}$ et du lemme 2.2.2. on déduit :

$$E_{N,\theta}^1 = \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^2 \gamma_2(t_i^N - t_{i-1}^N, \theta)$$

de sorte que :

$$\sup_{|\theta| \leq K} |E_{N,\theta}^1| \leq \delta_N \gamma_4(\delta_N, K)$$

où $\lim_{N \rightarrow \infty} \gamma_4(\delta_N, K) = \gamma_4(K)$.

Etudions maintenant $E_{N, \theta}^2$:

$$|E_{N, \theta}^2| \leq \frac{T}{2} p^{1/2} |\theta| (\text{tr} \{ [\frac{1}{T} \sum_{i=1}^{K_N} (\Pi_{t_i^N} - \Pi_{t_{i-1}^N})^{2\otimes} - \mathbb{1}_p]^2 \})^{1/2}.$$

Comme : $\mu_{\theta^*}^T - \lim_{N \rightarrow +\infty} \frac{1}{T} \sum_{i=1}^{K_N} (\Pi_{t_i^N} - \Pi_{t_{i-1}^N})^{2\otimes} = \mathbb{1}_p$,

on a :

$$\sup_{|\theta| \leq K} |E_{N, \theta}^2| \leq \frac{T}{2} p^{1/2} K \alpha(N)$$

où $\mu_{\theta^*}^T - \lim_{N \rightarrow +\infty} \alpha(N) = 0$ et, en vertu du lemme 2.2.3., $\{\frac{\alpha(N)}{\delta_N^{1/2}}\}$ est bornée en probabilité.

Grâce au lemme 2.2.2., on peut écrire :

$$E_{N, \theta}^3 = -\frac{1}{2} \text{tr} \left[\sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \Delta(t_i^N - t_{i-1}^N, \theta) (\Pi_{t_i^N} - \Pi_{t_{i-1}^N})^{2\otimes} \right]$$

et alors :

$$|E_{N, \theta}^3| \leq \delta_N |\Delta_1(\delta_N, \theta)| \text{tr} \left[\sum_{i=1}^{K_N} (\Pi_{t_i^N} - \Pi_{t_{i-1}^N})^{2\otimes} \right]$$

de sorte que l'on a :

$$\sup_{|\theta| \leq K} |E_{N, \theta}^3| \leq \delta_N \gamma_5(\delta_N, K) \beta(N)$$

où $\lim_{N \rightarrow +\infty} \gamma_5(\delta_N, K) = \gamma_5(K)$ et $\mu_{\theta^*}^T - \lim_{N \rightarrow +\infty} \beta(N) = pT$.

Maintenant, on écrit :

$$E_{N, \theta}^4 = -\frac{1}{2} \text{tr} \left[\sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^2 \Delta_2(t_i^N - t_{i-1}^N, \theta) \Pi_{t_{i-1}^N}^{2\otimes} \right]$$

et :

$$|E_{N, \theta}^4| \leq \delta_N \gamma_6(\delta_N, \theta) \text{tr} \left[\sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \Pi_{t_{i-1}^N}^{2\otimes} \right].$$

Alors :

$$\sup_{|\theta| \leq K} |E_{N, \theta}^4| \leq \delta_N \gamma_7(\delta_N, K) \zeta(N)$$

où $\lim_{N \rightarrow +\infty} \gamma_7(\delta_N, K) = \gamma_7(K)$ et $\mu_{\theta^*}^T - \lim_{N \rightarrow +\infty} \zeta(N) = \text{tr} \int_0^T \Pi_s^{2\otimes} ds$.

Pour achever la démonstration, on considère :

$$E_{N, \theta}^5 = - \text{tr} \left[\sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \Delta_3(t_i^N - t_{i-1}^N, \theta) \Pi_{t_{i-1}^N}^{t_i^N} (\Pi_{t_{i-1}^N} - \Pi_{t_i^N}) \right]$$

et :

$$\begin{aligned} |E_{N, \theta}^5| &\leq p^{1/2} \sum_{i=1}^{K_N} \left\{ \text{tr} \left[(t_i^N - t_{i-1}^N)^2 \Delta_3(t_i^N - t_{i-1}^N, \theta) \Pi_{t_{i-1}^N} \|\Pi_{t_i^N} - \Pi_{t_{i-1}^N}\|^2 \Pi_{t_{i-1}^N} \Delta_3(t_i^N - t_{i-1}^N, \theta) \right] \right\}^{1/2} \\ &\leq \delta_N^{1/2} \gamma_8(\delta_N, \theta) \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^{1/2} \|\Pi_{t_i^N} - \Pi_{t_{i-1}^N}\| \|\Pi_{t_{i-1}^N}\| \\ &\leq \delta_N^{1/2} \gamma_8(\delta_N, \theta) \left(\sum_{i=1}^{K_N} \|\Pi_{t_i^N} - \Pi_{t_{i-1}^N}\|^2 \right)^{1/2} \left(\sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \|\Pi_{t_{i-1}^N}\|^2 \right)^{1/2}. \end{aligned}$$

Aussi :

$$\sup_{|\theta| \leq K} |E_{N, \theta}^5| \leq \delta_N^{1/2} \gamma_8(\delta_N, K) \nu(N),$$

où $\lim_{N \rightarrow +\infty} \gamma_8(\delta_N, K) = \gamma_8(K)$, et $\mu_{\theta^*}^T - \lim_{N \rightarrow +\infty} \nu(N) = (pT)^{1/2} \left(\int_0^T \|\Pi_t\|^2 dt \right)^{1/2}$.

Nous étudions maintenant la différence $\tilde{L}_{N,T}^{(1)}(\theta, \cdot) - \hat{L}_{N,T}^{(1)}(\theta, \cdot)$ et démontrons l'assertion dans la proposition 2.2.1. pour $i=1$. On peut écrire :

$$\begin{aligned} (26) \quad &\tilde{L}_{N,T}^{(1)}(\theta) - \hat{L}_{N,T}^{(1)}(\theta) \\ &= -\frac{1}{2} \left\{ \frac{d}{d\theta} \left[\sum_{i=1}^{K_N} \text{Log det } \Sigma_{\theta, i}^N \right] - T \mathbb{1}_p \right\} \\ &\quad - \frac{1}{2} \left\{ \frac{d}{d\theta} \left[\text{tr} \sum_{i=1}^{K_N} (\Sigma_{\theta, i}^N)^{-1} (\Pi_{t_{i-1}^N} - \Pi_{t_i^N})^{2\otimes} + T \mathbb{1}_p \right] \right\} \end{aligned}$$

$$\begin{aligned}
 & -\frac{1}{2} \left\{ \frac{d}{d\theta} \left[\text{tr} \sum_{i=1}^{K_N} \left\{ (\mathbb{1}_p - e^{(t_i^N - t_{i-1}^N)\theta}) (\Sigma_{\theta,i}^N)^{-1} (\mathbb{1}_p - e^{(t_i^N - t_{i-1}^N)\theta})^{-t_{\theta\theta}(t_i^N - t_{i-1}^N)} \right\} \Pi_{N, t_{i-1}}^{2\otimes} \right] \right\} \\
 & - \frac{d}{d\theta} \left[\text{tr} \sum_{i=1}^{K_N} \left\{ (\Sigma_{\theta,i}^N)^{-1} (\mathbb{1}_p - e^{(t_i^N - t_{i-1}^N)\theta}) + \theta \right\} \Pi_{N, t_{i-1}}^t (\Pi_{N, t_i}^N - \Pi_{N, t_{i-1}}^N) \right].
 \end{aligned}$$

2.2.5. LEMME : "Pour toute constante $K > 0$, il existe une suite $\{\varepsilon_N^1(K) ; N \geq 1\}$ de variables aléatoires positives telle que, pour tout $\theta^* \in \mathfrak{L}(\mathbb{R}^p ; \mathbb{R}^p)$:

- (i) $\{\varepsilon_N^1(K) ; N \geq 1\}$ est bornée en $\mu_{\theta^*}^T$ -probabilité
- (ii) $\sup_{|\theta| \leq K} |\tilde{L}_{N,T}^{(1)}(\theta, \cdot) - \hat{L}_{N,T}^{(1)}(\theta, \cdot)| \leq \delta_N^{1/2} \varepsilon_N^1(K)$
 $\mu_{\theta^*}^T$ -presque sûrement.

Démonstration : $F_{N,\theta}^i$, $i = \overline{1,4}$ désignent les termes successifs de la somme dans (26). Calculons :

$$\begin{aligned}
 \frac{d}{d\theta_{k,\ell}} \sum_{i=1}^{K_N} \text{Log det } \Sigma_{\theta,i}^N &= \sum_{i=1}^{K_N} \frac{\delta_{k,\ell} (t_i^N - t_{i-1}^N) + (t_i^N - t_{i-1}^N)^2 \gamma_1^{(k,\ell)}(t_i^N - t_{i-1}^N, \theta)}{1 + (t_i^N - t_{i-1}^N) \text{tr } \theta + (t_i^N - t_{i-1}^N) \gamma_1(t_i^N - t_{i-1}^N, \theta)} \\
 &= T \delta_{k,\ell} + \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^2 \gamma_9^{k,\ell}(t_i^N - t_{i-1}^N, \theta)
 \end{aligned}$$

où $\sup_{|\theta| \leq K} |\gamma_9^{k,\ell}(h, \theta)| \leq \gamma_9^{k,\ell}(h, K)$ avec $\lim_{h \rightarrow 0} \gamma_9^{k,\ell}(h, K) = \gamma_9^{k,\ell}(K)$.

Alors :

$$\sup_{|\theta| \leq K} |F_{N,\theta}^1| \leq \delta_N \gamma_{10}(\delta_N, K)$$

où $\lim_{N \rightarrow +\infty} \gamma_{10}(\delta_N, K) = \gamma_{10}(K)$.

On évalue maintenant :

$$\begin{aligned} & \frac{d}{d\theta} \left\{ \text{tr} \sum_{i=1}^{K_N} (\Sigma_{\theta,i}^N)^{-1} (\Pi_{t_i^N - t_{i-1}^N}^N)^{2\otimes} \right\} \\ &= - \sum_{i=1}^{K_N} (\Pi_{t_i^N - t_{i-1}^N}^N)^{2\otimes} + \frac{d}{d\theta} \left\{ \text{tr} \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \Delta_4(t_i^N - t_{i-1}^N, \theta) (\Pi_{t_i^N - t_{i-1}^N}^N)^{2\otimes} \right\} \end{aligned}$$

et :

$$|F_{N,\theta}^2| \leq \frac{1}{2} |T \mathbb{1}_p - \sum_{i=1}^{K_N} (\Pi_{t_i^N - t_{i-1}^N}^N)^{2\otimes}| + \delta_N \gamma_{11}(\delta_N, \theta) \text{tr} \left\{ \sum_{i=1}^{K_N} (\Pi_{t_i^N - t_{i-1}^N}^N)^{2\otimes} \right\}$$

de sorte que, utilisant le lemme 2.2.3.,

$$\sup_{|\theta| \leq K} |F_{N,\theta}^2| \leq \delta_N^{1/2} V_N + \delta_N \gamma_{11}(\delta_N, K) \beta(N)$$

où $\{V_N\}$ est bornée en $\mu_{\theta^*}^T$ -probabilité, $\lim_{N \rightarrow +\infty} \gamma_{11}(\delta_N, K) = \gamma_{11}(K)$,

et $\mu_{\theta^*}^T - \lim_{N \rightarrow +\infty} \beta(N) = pT$.

Pour obtenir les inégalités relatives à $F_{N,\theta}^3$ et $F_{N,\theta}^4$ on suit la même démarche que celle utilisée pour $E_{N,\theta}^4$ et $E_{N,\theta}^5$ dans la démonstration du lemme 2.2.4. ■

A l'aide d'arguments similaires, on peut aisément démontrer que la différence $\tilde{L}_{N,T}^{(2)}(\theta, \cdot) - \hat{L}_{N,T}^{(2)}(\theta, \cdot)$ satisfait à une condition du même type et ainsi achever la démonstration de la proposition 2.2.1.

2.3. La fonction de vraisemblance approchée pour l'estimation basée sur un échantillonnage continu.

Nous étudions la différence $\hat{L}_{N,T}(\theta, \cdot) - L_T(\theta, \cdot)$. Ses propriétés sont formulées dans la proposition suivante :

2.3.1. PROPOSITION : "Pour toute constante $K > 0$ et tout $i = 0, 1, 2$, il existe une suite de variables aléatoires positives telle que, pour tout $\theta^* \in \mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$:

- (i) $\{\Delta_{N,T}^i(K) ; N \geq 1\}$ est bornée en $\mu_{\theta^*}^T$ -probabilité
- (jj) $\sup_{|\theta| \leq K} |\hat{L}_{N,T}^{(i)}(\theta, \cdot) - L_T^{(i)}(\theta, \cdot)| \leq \delta_N^{1/2} \Delta_{N,T}^i(K)$
 $\mu_{\theta^*}^T$ -presque sûrement."

La démonstration sera évidente au vu du lemme suivant :

2.3.2. LEMME : "On a :

- (i) $|\sum_{i=1}^{K_N} \Pi_N^{t_i} (\Pi_N^{t_i} - \Pi_N^{t_{i-1}}) - \int_0^T \Pi_s d t_{\Pi_s}| \leq \delta_N^{1/2} V_N$
- (ii) $|\sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \Pi_N^{2\otimes} - \int_0^T \Pi_s^{2\otimes} ds| \leq \delta_N^{1/2} W_N$

où $\{V_N\}_{N \geq 1}$ et $\{W_N\}_{N \geq 1}$ sont des suites de variables aléatoires bornées en $\mu_{\theta^*}^T$ -probabilité pour tout $\theta \in \mathfrak{L}(\mathbb{R}^p ; \mathbb{R}^p)$."

Démonstration : Posons :

$$\sum_{i=1}^{K_N} \Pi_N^{t_i} (\Pi_N^{t_i} - \Pi_N^{t_{i-1}}) - \int_0^T \Pi_s d t_{\Pi_s} = \delta_N^{1/2} \tilde{V}_N, \quad V_N = |\tilde{V}_N|$$

$$\sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \Pi_N^{2\otimes} - \int_0^T \Pi_s^{2\otimes} ds = \delta_N^{1/2} \tilde{W}_N \quad \text{et} \quad W_N = |\tilde{W}_N| .$$

Nous montrons que chaque terme dans \tilde{V}_N est borné en $\mu_{\theta^*}^T$ -probabilité. Supposons d'abord que $\theta = 0$, i.e. $\mu_{\theta^*}^T = \mu_{\theta^*}^T$; alors :

$$E_{\mu_{\theta^*}^T} \left| \sum_{i=1}^{K_N} \Pi_N^k (\Pi_N^{\ell} - \Pi_N^{\ell}) - \int_0^T \Pi_s^k d \Pi_s^{\ell} \right|^2$$

$$= E_{\mu_{\theta^*}^T} \left| \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\Pi_N^k - \Pi_s^k) d \Pi_s^{\ell} \right|^2$$

$$\leq 2 \sum_{i=1}^{K_N} E_{\mu_{\theta^*}^T} \left| \int_{t_{i-1}^N}^{t_i^N} (\Pi_N^k - \Pi_s^k) d \Pi_s^{\ell} \right|^2$$

$$= 2 \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (s-t_{i-1}^N) ds = \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N)^2 \leq \delta_N^T .$$

Maintenant, pour $\theta \in \mathfrak{L}(\mathbb{R}^p; \mathbb{R}^p)$, on peut écrire :

$$d\pi_t = \theta \pi_t dt + d\xi_t^\theta ; \quad t \in [0, T] ; \quad \pi_0 = 0$$

où $\{\xi_t^\theta\}$ est un processus de Wiener relativement à μ_θ^T . Alors,

$$\begin{aligned} & E_{\mu_\theta^T} \left| \sum_{i=1}^{K_N} \pi_{t_{i-1}^N}^k (\pi_{t_{i-1}^N}^\ell - \pi_{t_{i-1}^N}^\ell) - \int_0^T \pi_s^k d\pi_s^\ell \right|^2 \\ & \leq 4 E_{\mu_\theta^T} \left\{ \left| \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\xi_s^{\theta, k} - \xi_s^{\theta, \ell}) d\xi_s^{\theta, \ell} \right|^2 + \left| \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} \left(\int_{t_{i-1}^N}^s (\theta \pi_u)_k du \right) d\xi_s^{\theta, \ell} \right|^2 \right. \\ & \quad \left. + \left| \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\xi_s^{\theta, k} - \xi_s^{\theta, \ell}) (\theta \pi_s)_\ell ds \right|^2 + \left| \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} \left(\int_{t_{i-1}^N}^s (\theta \pi_u)_k du \right) (\theta \pi_s)_\ell ds \right|^2 \right\} \\ & \leq 8 \left\{ \delta_N^T + \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (s-t_{i-1}^N) \left[\int_{t_{i-1}^N}^s E_{\mu_\theta^T} |(\theta \pi_u)_\ell|^2 du \right] ds \right. \\ & \quad \left. + \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \int_{t_{i-1}^N}^{t_i^N} E_{\mu_\theta^T} \{ (\xi_s^{\theta, k} - \xi_s^{\theta, \ell})^2 (\theta \pi_s)_\ell^2 \} ds \right. \\ & \quad \left. + \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (s-t_{i-1}^N) E_{\mu_\theta^T} \left(\int_{t_{i-1}^N}^s (\theta \pi_u)_k^2 du \cdot (\theta \pi_s)_\ell^2 \right) ds \right\} . \end{aligned}$$

Utilisant le fait que les termes de la forme $E_{\mu_\theta^T}(\cdot)$ sont bornés sur $[0, T]$, l'assertion (i) du lemme peut aisément se déduire de cette majoration. On étudie maintenant \tilde{W}_N :

$$\begin{aligned}
 & E_{\mu_{\theta}^T} \left| \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \Pi_{t_{i-1}^N}^k \Pi_{t_{i-1}^N}^{\ell} - \int_0^T \Pi_s^k \Pi_s^{\ell} ds \right|^2 \\
 &= E_{\mu_{\theta}^T} \left| \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\Pi_{t_{i-1}^N}^k \Pi_{t_{i-1}^N}^{\ell} - \Pi_s^k \Pi_s^{\ell}) ds \right|^2 \\
 &\leq 2 \sum_{i=1}^{K_N} (t_i^N - t_{i-1}^N) \int_{t_{i-1}^N}^{t_i^N} E_{\mu_{\theta}^T} [|\Pi_{t_{i-1}^N}^k \Pi_{t_{i-1}^N}^{\ell} - \Pi_s^k \Pi_s^{\ell}|^2] ds .
 \end{aligned}$$

Cela fournit l'assertion (ii) puisque les termes $E_{\mu_{\theta}^T}(\cdot)$ sont bornés. ■

La démonstration de la proposition 2.3.1. est alors une simple application de ce lemme à l'étude de :

$$\hat{L}_{N,T}(\theta, \cdot) - L_T(\theta, \cdot) = \text{tr} \left\{ \theta \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\Pi_{t_{i-1}^N} - \Pi_s) d^t \Pi_s - \frac{1}{2} \theta \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\Pi_{t_{i-1}^N}^{2\otimes} - \Pi_s^{2\otimes}) ds \right\}$$

et de ses dérivées pour lesquelles on peut écrire :

$$\frac{d}{d\theta} (\hat{L}_{N,T}(\theta, \cdot) - L_T(\theta, \cdot)) = \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\Pi_{t_{i-1}^N} - \Pi_s) d^t \Pi_s - \theta \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\Pi_{t_{i-1}^N}^{2\otimes} - \Pi_s^{2\otimes}) ds$$

et, pour $u, v, k, \ell \in \{1, 2, \dots, p\}$:

$$\frac{\partial^2}{\partial \theta_{u,v} \partial \theta_{k,\ell}} (\hat{L}_{N,T}(\theta, \cdot) - L_T(\theta, \cdot)) = \delta_{u,k} \sum_{i=1}^{K_N} \int_{t_{i-1}^N}^{t_i^N} (\Pi_{t_{i-1}^N}^{\ell} \Pi_{t_{i-1}^N}^v - \Pi_s^{\ell} \Pi_s^v) ds .$$

2.3.3. Remarque : Pour les différences $\tilde{L}_{N,T}^{(i)}(\theta, \cdot) - L_T^{(i)}(\theta, \cdot)$ $i=0, 1, 2$, des propriétés analogues à celles de la proposition 2.3.1. pour $\hat{L}_{N,T}^{(i)}(\theta, \cdot) - L_T^{(i)}(\theta, \cdot)$, $i=0, 1, 2$, sont d'évidents corollaires des propositions 2.2.1. et 2.3.1.

2.4. Estimation du paramètre de dérive.

Nous pouvons démontrer la proposition suivante :

2.4.1. PROPOSITION : "Il existe une suite $\{\tilde{\theta}_{N,T}; N \geq 1\}$ d'estimateurs de θ telle que, pour tout $\theta^* \in \mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$ on a :

(j) $\tilde{\theta}_{N,T}$ est S_T^N -mesurable pour tout $N \geq 1$,

(jj) $\lim_{N \rightarrow +\infty} \mu_{\theta^*}^T[\tilde{L}_{N,T}(\tilde{\theta}_{N,T}, \cdot) = 0] = 1$,

(jjj) $\mu_{\theta^*}^T - \lim_{N \rightarrow +\infty} \tilde{\theta}_{N,T} = \hat{\theta}_T = [\int_0^T d\Pi_s \cdot t_{\Pi_s}] [\int_0^T \Pi_s^{2\otimes} ds]^{-1}$

(jv) les suites $\{\delta_N^{-1/2}(\tilde{\theta}_{N,T} - \hat{\theta}_T); N \geq 1\}$ et $\{\delta_N^{-1/2}(\hat{\theta}_{N,T} - \hat{\theta}_T); N \geq 1\}$ de variables aléatoires, où

$$\hat{\theta}_{N,T} = \left\{ \sum_{i=1}^{KN} (\Pi_{t_i^N} - \Pi_{t_{i-1}^N}) \cdot t_{\Pi_{t_{i-1}^N}} \right\} \left\{ \sum_{i=1}^{KN} (t_i^N - t_{i-1}^N) \Pi_{t_{i-1}^N}^{2\otimes} \right\}^{-1} ,$$

sont bornées en probabilité $\mu_{\theta^*}^T$."

Démonstration : Utilisant les propositions 2.2.1. et 2.3.1. et le fait que $\int_0^T \Pi_s^{2\otimes} ds$ est presque sûrement une matrice non singulière (proposition 1.3.1.), on peut aisément s'assurer que les hypothèses de la proposition I, 2.3.1. sont satisfaites pour $\{L_T, \tilde{L}_{N,T}; N \geq 1\}$ et $\{L_T, \hat{L}_{N,T}; N \geq 1\}$ avec $\gamma_N = \delta_N^{1/2}$.

Ainsi, la proposition I, 2.3.1. appliquée à $\{L_T, \tilde{L}_{N,T}; N \geq 1\}$ fournit les assertions (jj) et (jjj) et le fait que la suite $\{\delta_N^{-1/2}(\tilde{\theta}_{N,T} - \hat{\theta}_T); N \geq 1\}$ est bornée en probabilité $\mu_{\theta^*}^T$.

D'autre part, il est clair que $\hat{\theta}_{N,T}$ est l'unique solution de $\hat{L}_{N,T}(\theta, \cdot) = 0$ et donc $\{\hat{\theta}_{N,T}; N \geq 1\}$ n'est autre que la suite d'estimateurs fournie par la proposition I, 2.3.1. pour la suite $\{\hat{L}_{N,T}; N \geq 1\}$.

Enfin, compte-tenu des remarques I, 2.3.4, la suite $\{\tilde{\theta}_{N,T}; N \geq 1\}$ peut être choisie de telle manière que l'assertion (j) soit vérifiée. ■

2.4.2. Remarques :

a) Dans le cas où l'échantillonnage continu est possible et où le choix de l'estimateur $\hat{\theta}_{N,T}$ n'est alors imposé que par la nécessité d'approcher l'estimateur de maximum de vraisemblance $\hat{\theta}_T$, le résultat précédent fournit une base théorique pour ce calcul approché. En effet, si l'ordre de la variance (ou de la variance asymptotique $T \rightarrow +\infty$, Remarque 1.3.9.a) de l'estimateur exact est connu, sa comparaison avec l'ordre $\delta_N^{1/2}$ de l'erreur d'approximation peut permettre de juger de l'intérêt de raffiner la discrétisation.

b) D'autre part, si seul un échantillonnage discret est possible, on sait que $\hat{\theta}_{N,T}$ est "presque le meilleur estimateur", à l'ordre $\delta_N^{1/2}$ près, avec l'agréable propriété d'être facilement calculable.

c) Une approximation naturelle de la diffusion (X_t) solution de :

$$\begin{cases} dX_t = \theta X_t dt + G dW_t \\ X_0 = 0 \end{cases}$$

est fournie par une suite de variables aléatoires $(X_{t_1}, X_{t_2}, \dots, X_{t_N}, \dots)$ définie par :

$$\begin{cases} X_{t_{i+1}} - X_{t_i} = \theta X_{t_i} (t_{i+1} - t_i) + G(W_{t_{i+1}} - W_{t_i}) ; \quad i=0, 1, 2, \dots \\ t_0 = 0, \quad X_{t_0} = 0. \end{cases}$$

En particulier si le pas en temps est constant égal à h , posant $t_i = ih$ et $X_{t_i} = X_i$, le modèle s'écrit :

$$\begin{cases} X_{i+1} = (\mathbb{1}_p + \theta h) X_i + G \xi_i ; \quad i=0, 1, 2, \dots \\ X_0 = 0 \end{cases}$$

où $(\xi_0, \dots, \xi_N, \dots)$ est une suite de variables aléatoires gaussiennes indépendantes centrées de matrices de covariances $h.G.G^t$. C'est un

modèle autorégressif d'ordre 1 de série chronologique vectorielle. D'après les résultats classiques concernant ce modèle (voir par exemple T.W. ANDERSON [3] Ch. 5) l'estimateur de maximum de vraisemblance de $(\mathbb{1}_p + \theta h)$ est donné, pour des observations de $i = 0$ jusqu'à $i = N$, par :

$$(\widehat{\mathbb{1}_p + \theta h})_N = \frac{1}{N} \sum_{i=1}^N X_{i-1} \cdot {}^t X_{i-1} = \frac{1}{N} \sum_{i=1}^N X_i \cdot {}^t X_{i-1} .$$

L'estimateur $\hat{\theta}_N$ de θ , qui est alors donné par :

$$\begin{aligned} \hat{\theta}_N &= \left\{ - \sum_{i=1}^N X_{i-1} \cdot {}^t X_{i-1} + \sum_{i=1}^N X_i \cdot {}^t X_{i-1} \right\} \left\{ \sum_{i=1}^N h X_{i-1} \cdot {}^t X_{i-1} \right\}^{-1} \\ &= \left\{ \sum_{i=1}^N (X_i - X_{i-1}) \cdot {}^t X_{i-1} \right\} \left\{ \sum_{i=1}^N h X_{i-1} \cdot {}^t X_{i-1} \right\}^{-1} , \end{aligned}$$

coïncide avec l'estimateur $\hat{\theta}_{N, Nh}$ calculé pour la subdivision $\{0, h, 2h, \dots, Nh\}$ de $[0, Nh]$.

§ 3. - ESTIMATION DANS LES MODELES NON LINEAIRES A PARAMETRES MULTIPLICATIFS DANS LES COEFFICIENTS DE DIFFUSION ET DE DERIVE.

On s'intéresse ici à un modèle de la forme :

$$(27) \quad \begin{cases} dX_t = [a_0(t, X) + \sum_{i=1}^r \varphi_i a_i(t, X)] dt + [\sum_{i=r+1}^d \varphi_i \sigma_i(t, X)] dW_t ; & t \geq 0 \\ X_0 = 0 \end{cases}$$

où : - $\varphi = (\varphi_{\Delta}, \varphi_{\bar{\Delta}}) = ((\varphi_1, \dots, \varphi_r), (\varphi_{r+1}, \dots, \varphi_d))$ est un paramètre à estimer dans \mathbb{R}^d ,

- a_i , $i = \overline{0, r}$ sont des applications données de $\mathbb{R}_+ \times C$ dans \mathbb{R}^p et σ_i , $i = \overline{r+1, d}$ sont des applications données de $\mathbb{R}_+ \times C$ dans $\mathfrak{L}(\mathbb{R}^p; \mathbb{R}^p)$ vérifiant des conditions (voir I. § 1) assurant que les hypothèses (D1) et (D2) du I. § 1.4. sont satisfaites.

Alors on a :

$$\frac{dP_\varphi}{dQ_\varphi} / S_T = \exp\left\{ \int_0^T \langle \sigma_{\varphi_\Delta}^{-1}(t, \cdot) a_{\varphi_\Delta}(t, \cdot), \sigma_{\varphi_\Delta}^{-1}(t, \cdot) d\Pi_t \rangle - \frac{1}{2} \int_0^T \|\sigma_{\varphi_\Delta}^{-1}(t, \cdot) a_{\varphi_\Delta}(t, \cdot)\|^2 dt \right\}$$

- où :
- $a_{\varphi_\Delta}(t, x) = a_0(t, x) + \sum_{i=1}^r \varphi_i a_i(t, x)$
 - $\sigma_{\varphi_\Delta}(t, x) = \sum_{i=r+1}^d \varphi_i \sigma_i(t, x) \quad ; (t, x) \in \mathbb{R}_+ \times C$
 - P_φ (resp. Q_{φ_Δ}) est la solution faible de (27) (resp. $dX_t = \sigma_{\varphi_\Delta}(t, \cdot) dW_t \quad ; X_0 = 0, \quad t \geq 0$)
 - l'intégrale stochastique est prise relativement à Q_{φ_Δ} .

3.1. Estimation des paramètres du coefficient de diffusion.

Afin de pouvoir utiliser la méthode du maximum de vraisemblance pour estimer le paramètre φ_Δ intervenant dans le coefficient de dérive, il convient que le paramètre φ_Δ dont dépend le coefficient de diffusion soit connu. Nous montrons donc que sous certaines conditions φ_Δ est identifiable sur un intervalle de temps fini : on considère, pour $t \in]0, T]$ et $N \geq 1$, la statistique $\hat{\sigma}_t^N$ définie sur la structure statistique $(C, S_T, \{P_\varphi^T ; \varphi \in \mathbb{R}^d\})$ par :

$$\hat{\sigma}_t^N = \sum_{i=1}^{2N} \left(\Pi_{i \frac{t}{2N}} - \Pi_{(i-1) \frac{t}{2N}} \right)^{2\otimes}$$

On montre que la suite de statistiques $\{\hat{\sigma}_t^N ; N \geq 1\}$ est convergente en probabilité :

3.1.1. LEMME : "Pour tout $\varphi \in \mathbb{R}^d$ et tout $t \in]0, T]$ on a :

$$P_\varphi^T - \lim_{N \rightarrow +\infty} \hat{\sigma}_t^N = \int_0^t \sigma_{\varphi_\Delta}(s, \cdot) \cdot {}^t \sigma_{\varphi_\Delta}(s, \cdot) ds ."$$

Démonstration : Relativement à P_φ , le processus

$$(V_t)_{t \geq 0} = (\Pi_t - X_t)_{t \geq 0} \quad \text{où} \quad X_t = \int_0^t a_{\varphi_\Delta}(s, \cdot) ds ; \quad t \geq 0,$$

est une martingale continue de carré intégrable par rapport à $(S_t)_{t \geq 0}$, qui admet la représentation :

$$V_t = \int_0^t \sigma_{\varphi_\Delta}(s, \cdot) d\tilde{\Pi}_s ; \quad t \geq 0,$$

où $(\tilde{\Pi}_t)_{t \geq 0}$ est un mouvement brownien. Alors le processus naturel de $V_t \otimes V_t$ est :

$$\begin{aligned} \langle V \rangle_t &= \int_0^t \sigma_{\varphi_\Delta}(s, \cdot) \cdot {}^t\sigma_{\varphi_\Delta}(s, \cdot) d\langle \tilde{\Pi} \rangle_s \\ &= \int_0^t \sigma_{\varphi_\Delta}(s, \cdot) \cdot {}^t\sigma_{\varphi_\Delta}(s, \cdot) ds \quad P_\varphi \text{ presque sûrement.} \end{aligned}$$

On a donc :

$$\lim_{N \rightarrow +\infty} \sum_{i=1}^{2^N} (V_{i \frac{t}{2^N}} - V_{(i-1) \frac{t}{2^N}})^{2\otimes} = \int_0^t \sigma_{\varphi_\Delta}(s, \cdot) \cdot {}^t\sigma_{\varphi_\Delta}(s, \cdot) ds$$

au sens de la convergence en moyenne relativement à P_φ^T .

D'autre part, le processus $(X_s)_{s \geq 0}$ est à trajectoires à variation bornée sur l'intervalle $[0, t]$ donc :

$$\lim_{N \rightarrow +\infty} \sum_{i=1}^{2^N} \psi(t_i^N) \cdot (X_{i \frac{t}{2^N}} - X_{(i-1) \frac{t}{2^N}}) = \int_0^t \psi(s) d^t X_s,$$

au sens de la convergence P_φ^T presque sûre, si $(\psi(s))_{s \in [0, t]}$ est un processus à trajectoires continues et $t_i^N \in [(i-1) \frac{t}{2^N}, i \frac{t}{2^N}]$, $i=1, 2^N$, $N \in \mathbb{N}^*$.

On en déduit immédiatement le résultat du lemme. ■

Nous pouvons alors énoncer :

3.1.2. PROPOSITION : "S'il existe un instant $t > 0$ tel que pour tout $i \in \{r+1, \dots, d\}$ il existe $(k, \ell) \in \{1, \dots, p\}^2$ tel que $\int_0^t (\sigma_i(s, \cdot) \cdot {}^t\sigma_i(s, \cdot))_{k, \ell} ds \neq 0$ et, pour tout $j \in \{r+1, \dots, d\} \setminus \{i\}$,

$\int_0^t (\sigma_i(s, \cdot) \cdot {}^t \sigma_j(s, \cdot))_{k, \ell} ds = 0$, alors le paramètre $(\varphi_{r+1}^2, \dots, \varphi_d^2)$ est
fortement identifiable en probabilité sur l'intervalle de temps $[0, t]$;
de manière précise on a, pour $i \in \{r+1, \dots, d\}$,

$$P_{\varphi} - \lim_{N \rightarrow +\infty} \frac{(\hat{\sigma}_t^N)_{k, \ell}}{\int_0^t (\sigma_i(s, \cdot) \cdot {}^t \sigma_i(s, \cdot))_{k, \ell} ds} = \varphi_i^2$$

où (k, ℓ) est défini ci-dessus."

Démonstration : D'après le lemme 3.1.1. on a :

$$P_{\varphi} - \lim_{N \rightarrow +\infty} \hat{\sigma}_t^N = \sum_{i, j=r+1}^d \varphi_i \varphi_j \int_0^t \sigma_i(s, \cdot) \cdot {}^t \sigma_j(s, \cdot) ds$$

soit, compte-tenu de l'hypothèse :

$$P_{\varphi} - \lim_{N \rightarrow +\infty} (\hat{\sigma}_t^N)_{k, \ell} = \varphi_i^2 \int_0^t (\sigma_i(s, \cdot) \cdot {}^t \sigma_i(s, \cdot))_{k, \ell} ds .$$

Puisque $\int_0^t (\sigma_i(s, \cdot) \cdot {}^t \sigma_i(s, \cdot))_{k, \ell} \neq 0$, le résultat est acquis. ■

3.1.3. Remarques.

a) Dans le chapitre suivant nous envisagerons d'autres hypothèses, que celles de la proposition 3.1.2., assurant l'identifiabilité du paramètre φ_{Δ} .

b) Les hypothèses de la proposition 3.1.2. sont par exemple vérifiées si $r = d - p$ et $\sigma_{\varphi_{\Delta}}(s, x)$ est diagonale avec :

$$(\sigma_{\varphi_{\Delta}}(s, x))_{i, j} = \begin{cases} \varphi_i (\sigma_i(s, x))_{ii} & \text{si } i=j \\ 0 & \text{si } i \neq j \end{cases} \quad (i, j) \in \{d-p+1, \dots, d\}^2$$

et $(\sigma_i(s, x))_{i, i} \neq 0$, $i \in \{d-p+1, \dots, d\}$.

En particulier, si $p = 1$, φ_d^2 est fortement identifiable en probabilité sur $[0, t]$ par la suite de statistiques $\{\hat{\sigma}_t^N / \int_0^t \sigma_d^2(s, \cdot) ds ; N \geq 1\}$.

c) Si $\sigma_{\varphi_{\bar{\Delta}}}$ ne dépend pas de (t, x) , alors, compte-tenu du lemme 1.2.2., $\sigma_{\varphi_{\bar{\Delta}}} \cdot \sigma_{\varphi_{\bar{\Sigma}}}$ est fortement identifiable presque sûrement sur tout intervalle de temps fini $[0, t]$, par la suite de statistiques $\{\frac{1}{t} \hat{\sigma}_t^N ; N \geq 1\}$.

d) Si certains paramètres du modèle (27) apparaissent simultanément dans le coefficient de diffusion et dans le coefficient de dérive i.e. si pour $i \in I \subset \Delta = \{1, \dots, r\}$, il existe $j \in \bar{\Delta} = \{r+1, \dots, d\}$ tel que $\alpha_i = \alpha_j$, le problème d'estimation des paramètres de dérive, après estimation des paramètres de diffusion, est un problème relatif à un paramètre vectoriel de dimension $r \setminus |I|$.

3.2. Estimation des paramètres de dérive par maximum de vraisemblance.

On considère un cas particulier du modèle (27) s'écrivant :

$$(28) \quad \begin{cases} dX_t = [a_0(t, X) + \sum_{i=1}^d \varphi_i a_i(t, X)] dt + \sigma(t, X) dW_t & ; t \geq 0 \\ X_0 = 0 \end{cases}$$

où $\sigma(\dots)$ ne dépend pas de paramètres inconnus éventuellement parce qu'ils auraient été identifiés sur un intervalle de temps fini.

La structure statistique $(C, S_T, \{P_\varphi^T ; \varphi \in \mathbb{R}^d\})$ est alors dominée avec :

$$\frac{dP_\varphi}{dQ} /_{S_T} = \exp\left\{ \int_0^T \langle \sigma^{-1}(s, \cdot) a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) d\Pi_s \rangle - \frac{1}{2} \int_0^T \|\sigma^{-1}(s, \cdot) a_\varphi(s, \cdot)\|^2 ds \right\}$$

- où :
- $\varphi = (\varphi_1, \dots, \varphi_d) \in \mathbb{R}^d$
 - $a_\varphi(s, x) = a_0(t, x) + \sum_{i=1}^d \varphi_i a_i(t, x) ; (s, x) \in \mathbb{R}_+ \times C$
 - P_φ (resp. Q) est la solution faible de (28) (resp. $dX_t = \sigma(t, X) dW_t$, $X_0 = 0 ; t \geq 0$)
 - l'intégrale stochastique est prise relativement à Q .

La fonction de Log-vraisemblance $L_T(\varphi, .)$ pour l'estimation de σ au vu d'une observation sur $[0, T]$ s'écrit :

$$\begin{aligned}
 L_T(\varphi, .) &= \int_0^T \langle \sigma^{-1}(t, .) a_0(t, .), \sigma^{-1}(t, .) d\Pi_t \rangle \\
 &+ \sum_{i=1}^d \varphi_i \int_0^T \langle \sigma^{-1}(t, .) a_i(t, .), \sigma^{-1}(t, .) d\Pi_t \rangle \\
 &- \frac{1}{2} \int_0^T \langle (\sigma(t, .) \cdot {}^t\sigma(t, .))^{-1} a_0(t, .), a_0(t, .) \rangle dt \\
 &- \sum_{i=1}^d \varphi_i \int_0^T \langle (\sigma(t, .) \cdot {}^t\sigma(t, .))^{-1} a_0(t, .), a_i(t, .) \rangle dt \\
 &- \frac{1}{2} \sum_{i,j=1,d} \varphi_i \varphi_j \int_0^T \langle (\sigma(t, .) \cdot {}^t\sigma(t, .))^{-1} a_i(t, .), a_j(t, .) \rangle dt .
 \end{aligned}$$

Sa dérivée première $L_T^{(1)}(\varphi, .)$ est donnée par :

$$\begin{aligned}
 \frac{\partial L_T}{\partial \varphi_k}(\varphi, .) &= \int_0^T \langle (\sigma(t, .) \cdot {}^t\sigma(t, .))^{-1} a_k(t, .), d\Pi_t - a_0(t, .) \rangle \\
 &- \sum_{i=1}^d \varphi_i \int_0^T \langle (\sigma(t, .) \cdot {}^t\sigma(t, .))^{-1} a_k(t, .), a_i(t, .) \rangle dt \\
 k &= \overline{1, d} .
 \end{aligned}$$

De même sa dérivée seconde $L_T^{(2)}(\varphi, .)$ est définie par :

$$\begin{aligned}
 \frac{\partial^2 L_T}{\partial \varphi_\ell \partial \varphi_k}(\varphi, .) &= - \int_0^T \langle (\sigma(t, .) \cdot {}^t\sigma(t, .))^{-1} a_k(t, .), a_\ell(t, .) \rangle dt \\
 k, \ell &= \overline{1, d} .
 \end{aligned}$$

L'équation de vraisemblance $L_T^{(1)}(\varphi, .) = 0$ s'écrit sous la forme :

$$(29) \quad A_T \varphi = B_T$$

où A_T est la matrice opposée de la matrice des dérivées partielles secondes de $L_T(\varphi, .)$ (qui ne dépendent pas de φ) et B_T est le vecteur de dimension d de composantes :

$$\int_0^T \langle (\sigma(t, .) \cdot {}^t\sigma(t, .))^{-1} a_k(t, .), d\Pi_t - a_0(t, .) \rangle ; \quad k = \overline{1, d} .$$

On a alors l'énoncé suivant qui recouvre un résultat de A.F. TARASKIN [49] :

3.2.1. PROPOSITION : " φ^* désignant la vraie valeur du paramètre φ , si les conditions suivantes sont satisfaites :

$$(i) \quad \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \sigma^{-1}(t, \cdot) A(t, \cdot) \cdot {}^t A(t, \cdot) \cdot {}^t \sigma^{-1}(t, \cdot) dt = \Gamma \quad ,$$

P_{φ^*} presque sûrement

$$(ii) \quad \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T E_{P_{\varphi^*}} [\sigma^{-1}(t, \cdot) A(t, \cdot) \cdot {}^t A(t, \cdot) \cdot {}^t \sigma^{-1}(t, \cdot)] dt = \Gamma$$

où $A(t, \cdot)$ est la matrice dont les colonnes sont $a_1(t, \cdot), \dots, a_d(t, \cdot)$ et
" est une matrice non singulière, alors dans le modèle (28) φ^* est
asymptotiquement identifiable en probabilité par une famille $\{\hat{\varphi}_T; T > 0\}$
d'estimateurs telle que :

(j) pour tout $T > 0$, $\hat{\varphi}_T$ est S_T -mesurable,

(jj) $\lim_{T \rightarrow +\infty} P_{\varphi^*} [A_T \hat{\varphi}_T = B_T] = 1$,

(jjj) $\mathcal{L}_{P_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{1/2} (\hat{\varphi}_T - \varphi^*) = N(0, \Gamma^{-1})$."

Démonstration : Elle consiste à vérifier que les hypothèses de la proposition I. 2.2.1. sont vérifiées. On a :

$$(30) \quad L_{T,k}^{(1)}(\varphi^*, \cdot) = \int_0^T \langle \sigma^{-1}(t, \cdot) a_k(t, \cdot), \sigma^{-1}(t, \cdot) (d\Pi_t - a_{\varphi^*}(t, \cdot) dt) \rangle$$

où, relativement à P_{φ^*} , le processus $(\int_0^t \sigma^{-1}(s, \cdot) (d\Pi_s - a_{\varphi^*}(s, \cdot) ds) ; t \geq 0)$ est un processus de Wiener standard.

Alors $E_{P_{\varphi^*}} |L_{T,k}^{(1)}(\varphi^*, \cdot)|^2 \leq d \int_0^T E_{P_{\varphi^*}} \|\sigma^{-1}(t, \cdot) a_k(t, \cdot)\|^2 dt$

et d'après l'hypothèse (ii) on a :

$$P_{\varphi^*} - \lim_{T \rightarrow +\infty} \frac{1}{T} L_{T,k}^{(1)}(\varphi^*, \cdot) = 0 \quad .$$

L'hypothèse (E1) de la proposition I. 2.2.1. est donc vérifiée.

L'hypothèse (E2) de cette proposition est une conséquence immédiate de l'hypothèse (i),

l'hypothèse (E4) étant ici trivialement vérifiée puisque $L_T^{(2)}(\varphi, \cdot)$ ne dépend pas de φ .

Il reste donc à vérifier l'hypothèse (E3). Or d'après (30), (i) et (ii), le théorème I. 3.2.1. assure que :

$$\mathfrak{L}_{P_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2} L_T(\varphi^*, \cdot) = N(0, \Gamma) .$$

La proposition I. 2.2.1. fournit alors les conclusions (j) à (jjj). ■

On indique maintenant un résultat de A.F. TARASKIN [49] à propos d'un modèle particulier de la forme (28) :

3.2.2. THEOREME [49] : "On suppose que les applications a_i , $i=1, d$ et σ sont telles que :

$$\begin{aligned} a_i(t, x) &\equiv a_i(t) \\ &(t, x) \in \mathbb{R}_+ \times C \\ \text{et } \tau(t, x) &\equiv \sigma(t) \end{aligned}$$

et que les applications a_1, \dots, a_d sont linéairement indépendantes.

Alors la solution $\hat{\varphi}_T$ de l'équation de vraisemblance (29) est un estimateur sans biais et efficace de φ^* . De plus le vecteur aléatoire $[\hat{\varphi}_T - \varphi^*]$ a une distribution gaussienne (centrée et) de matrice de covariances $-A_T^{-1}$."

Démonstration [49] : Compte-tenu de l'hypothèse d'indépendance de a_1, \dots, a_d , la matrice A_T est inversible et l'équation (29) a une solution unique :

$$\hat{\varphi}_T = A_T^{-1} B_T .$$

On a :

$$L_T^{(1)}(\varphi^*, \cdot) = B_T - A_T \varphi^*$$

$$0 = B_T - A_T \hat{\varphi}_T$$

d'où l'on déduit que :

$$L_T^{(1)}(\varphi^*, \dots) = A_T(\hat{\varphi}_T - \varphi^*) .$$

Or, compte-tenu de (30), pour $k = \overline{1, d}$, $E_{P_{\varphi^*}}(L_{T,k}^{(1)}(\varphi^*, \dots)) = 0$, donc on a :

$$E_{P_{\varphi^*}}(\hat{\varphi}_T - \varphi^*) = A_T^{-1} E_{P_{\varphi^*}}(L_T^{(1)}(\varphi^*, \dots)) = 0$$

et $\hat{\varphi}_T$ est un estimateur sans biais.

D'après le théorème I. 3.2.1., le vecteur $L_T^{(1)}(\varphi^*, \dots)$ est, relativement à P_{φ^*} , gaussien centré de matrice de covariances $-A_T$. Alors $\hat{\varphi}_T - \varphi^*$ est de loi $N(0, -A_T^{-1})$.

D'après l'inégalité de Cramer-Rao, pour tout estimateur $\tilde{\varphi}_T$ sans biais de φ^* on a $E((\tilde{\varphi}_T - \varphi^*) \cdot {}^t(\tilde{\varphi}_T - \varphi^*)) \geq -A_T^{-1}$. Comme l'égalité est vérifiée pour $\hat{\varphi}_T$, $\hat{\varphi}_T$ est un estimateur efficace. ■

3.2.3. Remarques.

a) Supposons que dans le modèle scalaire :

$$dX_t = a_0(t, X)dt + \phi a_1(t)dt + \sigma(t)dW_t$$

le coefficient $a_1(t)$ représente un terme de contrôle à la disposition de l'observateur. Un problème naturel est alors de choisir le contrôle de manière à maximiser la qualité de l'estimateur :

$$\hat{a}_T^{a_1} = \left[\int_0^T \frac{a_1^2(t)}{\sigma^2(t)} (d\tilde{\Gamma}_t - a_0(t, \cdot)dt) \right] \left[\int_0^T \frac{a_1^2(t)}{\sigma^2(t)} dt \right]^{-1}$$

dont on sait par le théorème 3.2.2. qu'il est sans biais et efficace, soit

de variance $\left[\int_0^T \frac{a_1^2(t)}{\sigma^2(t)} dt \right]^{-1}$. Maximiser la qualité de $\hat{a}_T^{a_1}$ revient à

minimiser sa variance soit à maximiser $\int_0^T \frac{a_1^2(t)}{\sigma^2(t)} dt$. Il est clair que le

meilleur contrôle pour l'estimation de φ , dans le cas où les contrôles admissibles sont à valeurs dans un intervalle fermé borné $[u_1, u_2]$, est donné par :

$$a_1(t) = \begin{cases} u_2 & \text{si } |u_2| \geq |u_1| \\ u_1 & \text{si } |u_1| \geq |u_2| . \end{cases}$$

Dans cet esprit, LOPEZ TOLEDO [37] traite du problème du choix optimal d'une entrée ("input") pour l'identification de systèmes stochastiques linéaires.

b) Contrairement à ce qui se produit dans un modèle du type de celui envisagé dans le théorème 3.2.2., en général, lorsque les coefficients de l'équation régissant la diffusion dépendent à la fois de t et de x , même si l'équation de vraisemblance admet une solution unique, l'estimateur de maximum de vraisemblance n'est pas sans biais. Ainsi, reprenant le cas linéaire unidimensionnel :

$$dX_t = \theta X_t dt + dW_t ,$$

on a :

$$\int_0^T \Pi_s d\Pi_s = \hat{\theta}_T \int_0^T \Pi_s^2 ds$$

ou :

$$\int_0^T \Pi_s (d\Pi_s - \theta^* \Pi_s ds) = (\hat{\theta}_T - \theta^*) \int_0^T \Pi_s^2 ds$$

où, relativement à P_{θ^*} , le processus $(\Pi_t - \int_0^t \theta^* \Pi_s ds ; t \geq 0)$ est un processus de Wiener. Alors, on a :

$$E_{P_{\theta^*}} \left\{ \int_0^T \Pi_s (d\Pi_s - \theta^* \Pi_s ds) \right\} = 0$$

alors qu'il faudrait, pour que $\hat{\theta}_T$ soit un estimateur sans biais, que :

$$E_{P_{\theta^*}} \left\{ \left[\int_0^T \Pi_s (d\Pi_s - \theta^* \Pi_s ds) \right] \left[\int_0^T \Pi_s^2 ds \right]^{-1} \right\} = 0 .$$

Dans le paragraphe suivant nous envisageons, à l'aide d'une technique proposée par R.S. LIPCER et A.N. SHIRYAYEV, la construction d'estimateurs sans biais et gaussiens à partir des estimateurs de maximum de vraisemblance.

3.3. Estimation "séquentielle" de maximum de vraisemblance des paramètres de dérive.

On considère un cas particulier du modèle (28) :

$$(31) \quad \begin{aligned} dX_t &= [a_0(t, X) + \varphi a_1(t, X)] dt + \sigma(t, X) dW_t \quad ; \quad t \geq 0 \\ X_0 &= 0 \end{aligned}$$

où φ est un paramètre scalaire.

D'après l'étude précédente, l'estimateur de maximum de vraisemblance $\hat{\varphi}_T$ de φ est donné par l'équation (29) i.e. ici par :

$$[\int_0^T \|\sigma^{-1}(t, \cdot) a_1(t, \cdot)\|^2 dt] \cdot \hat{\varphi}_T = \int_0^T \langle (\sigma(t, \cdot) \cdot {}^t \sigma(t, \cdot))^{-1} a_1(t, \cdot), d\Pi_t - a_0(t, \cdot) dt \rangle$$

soit :

$$\hat{\varphi}_T = \frac{\int_0^T \langle (\sigma(t, \cdot) \cdot {}^t \sigma(t, \cdot))^{-1} a_1(t, \cdot), d\Pi_t - a_0(t, \cdot) dt \rangle}{\int_0^T \|\sigma^{-1}(t, \cdot) a_1(t, \cdot)\|^2 dt}$$

que l'on peut encore écrire sous la forme :

$$(32) \quad \hat{\varphi}_T = \varphi + \frac{\int_0^T \langle \sigma^{-1}(t, \cdot) a_1(t, \cdot), \sigma^{-1}(t, \cdot) [d\Pi_t - (a_0(t, \cdot) + \varphi a_1(t, \cdot))] dt \rangle}{\int_0^T \|\sigma^{-1}(t, \cdot) a_1(t, \cdot)\|^2 dt}$$

où, relativement à P_φ , le processus $(\int_0^t \sigma^{-1}(s, \cdot) [d\Pi_s - (a_0(s, \cdot) + \varphi a_1(s, \cdot))] ds ; t \geq 0)$ est un processus de Wiener.

Sous des hypothèses convenables (par exemple $0 < c \leq \|\sigma^{-1}(t, \cdot) a_1(t, \cdot)\| \leq C < +\infty$) on peut montrer, comme R.S. LIPCER et A.N. SHIRYAYEV l'ont fait dans [36] Ch. 17, dans le cas scalaire, que

$$E_{P_\varphi} (\hat{\varphi}_T - \varphi) = \frac{\partial}{\partial \varphi} E_{P_\varphi} [(\int_0^T \|\sigma^{-1}(t, \cdot) a_1(t, \cdot)\|^2 dt)^{-1}] .$$

En général l'estimateur $\hat{\varphi}_T$ est donc biaisé. Cependant, suivant une idée des auteurs déjà cités, on peut démontrer la proposition suivante :

3.3.1. PROPOSITION : "On suppose que

$P_\varphi [\int_0^{+\infty} \|\sigma^{-1}(t, \cdot) a_1(t, \cdot)\|^2 dt = +\infty] = 1$ et on définit, pour $H \geq 0$, le
temps d'arrêt τ_H par :

$$\tau_H = \inf\{t \geq 0 : \int_0^t \|\sigma^{-1}(s, \cdot) a_1(s, \cdot)\|^2 ds \geq H\} .$$

Alors l'estimateur de maximum de vraisemblance séquentiel

$$\tilde{\varphi}_H = \hat{\varphi}_{\tau_H} = \frac{1}{H} \int_0^{\tau_H} \langle (\sigma(t, \cdot) \cdot {}^t\sigma(t, \cdot))^{-1} a_1(t) , d\Pi_t - a_0(t, \cdot) dt \rangle$$

de φ dans le modèle (31) est un estimateur sans biais, distribué sui-
vant une loi gaussienne de variance $\frac{1}{H}$."

Démonstration : D'après un résultat classique sur le change-
ment de temps dans une martingale (voir par exemple R. PRIOURET [43]),
l'hypothèse $P_\varphi [\int_0^{+\infty} \|\sigma^{-1}(s, \cdot) a_1(s, \cdot)\|^2 ds = +\infty] = 1$ assure que le proces-
sus :

$$\left(\int_0^{\tau_H} \langle \sigma^{-1}(t, \cdot) a_1(t, \cdot) , \sigma^{-1}(t, \cdot) [d\Pi_t - (a_0(t, \cdot) + \varphi a_1(t, \cdot)) dt] \rangle ; H \geq 0 \right)$$

est un mouvement brownien standard $(B_H ; H \geq 0)$ relativement à P_φ .
Alors, d'après (32), on a :

$$\tilde{\varphi}_H = \hat{\varphi}_{\tau_H} = \varphi + \frac{1}{H} B_H .$$

On en déduit immédiatement les conclusions de la proposition. ■

3.3.2. Remarques.

a) On peut penser à utiliser cette technique de changement de temps dans le cas d'un paramètre à estimer qui soit vectoriel. Considérons le cas linéaire vectoriel :

$$dX_t = \theta X_t dt + dW_t .$$

On a alors :

$$\int_0^T d\Pi_s \cdot {}^t\Pi_s = \hat{\theta}_T \int_0^T \Pi_s \cdot {}^t\Pi_s ds$$

ou
$$\int_0^T (d\Pi_s - \theta \Pi_s ds) \cdot {}^t\Pi_s = (\hat{\theta}_T - \theta) \int_0^T \Pi_s \cdot {}^t\Pi_s ds$$

où $(\Pi_t^\theta)_{t \geq 0}$ défini par $\Pi_t^\theta = \Pi_t - \int_0^t \theta \Pi_s ds$ est un processus de Wiener relativement à P_θ .

On a aussi :

$$\left(\int_0^T \Pi_s \cdot {}^t\Pi_s ds \right) ({}^{t\hat{\theta}_T - t\theta})_{.i} = \begin{pmatrix} \int_0^T \Pi_s^1 \cdot d\Pi_s^{\theta,i} \\ \vdots \\ \int_0^T \Pi_s^p \cdot d\Pi_s^{\theta,i} \end{pmatrix} = M_T^i, \quad i = \overline{1, p},$$

où $({}^{t\hat{\theta}_T - t\theta})_{.i}$ désigne la i ème colonne de ${}^{t\hat{\theta}_T - t\theta}$ et Π_s^i (resp. $\Pi_s^{\theta,i}$) est la i ème composante de Π_s (resp. Π_s^θ). Remarquons que le processus $(\int_0^T \Pi_s \cdot {}^t\Pi_s ds ; T \geq 0)$ n'est autre que le processus naturel de $(M_T^i \otimes M_T^i ; T \geq 0)$.

Il s'agirait ici d'introduire non pas un temps d'arrêt mais les p temps d'arrêt :

$$\tau_H^j = \inf\{t \geq 0 : \int_0^t (\Pi_s^j)^2 ds \geq H\} ; \quad j = \overline{1, p},$$

tels que $(\int_0^{\tau_H^j} \Pi_s^j d\Pi_s^{\theta,i} ; H \geq 0)$, $j = \overline{1, p}$, soient des mouvements browniens.

Mais cela ne permet pas les mêmes développements que ceux conduits dans le cas scalaire puisqu'ici, dans chaque ligne $\int_0^{\tau_H^j} \Pi_s^j \cdot {}^t\Pi_s ds$, seul le terme de la j ème colonne est constant égale à H .

b) Toutefois dans [41], A.A. NOVIKOV considère le cas particulier suivant pour $p = 2$, $X_t = \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix}$, $W_t = \begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix}$:

$$\begin{cases} dX_t^1 = [-\varphi_1 u(t, X) - \varphi_2 v(t, X)] dt + dW_t^1 \\ dX_t^2 = [\varphi_2 u(t, X) - \varphi_1 v(t, X)] dt + dW_t^2 \end{cases}$$

Le modèle peut s'écrire sous la forme (28) :

$$dX_t = [\varphi_1 a_1(t, X) + \varphi_2 a_2(t, X)] dt + dW_t$$

en posant $a_1(t, x) = \begin{pmatrix} -u(t, x) \\ -v(t, x) \end{pmatrix}$, $a_2(t, x) = \begin{pmatrix} -v(t, x) \\ u(t, x) \end{pmatrix}$, $(t, x) \in \mathbb{R}_+ \times C$.

Alors, compte-tenu de ce que $\langle a_1(t, x), a_2(t, x) \rangle = 0$ et $\|a_1(t, x)\|^2 = \|a_2(t, x)\|^2 = u^2(t, x) + v^2(t, x)$, la matrice A_T de l'équation de vraisemblance (29) est diagonale :

$$A_T = \begin{pmatrix} \int_0^T [u^2(t, \cdot) + v^2(t, \cdot)] dt & 0 \\ 0 & \int_0^T [u^2(t, \cdot) + v^2(t, \cdot)] dt \end{pmatrix}$$

Les estimateurs de maximum de vraisemblance de φ_1 et φ_2 sont donnés par :

$$\hat{\varphi}_{1, T} = -\left(\int_0^T u(t, \cdot) d\bar{\Pi}_t^1 + \int_0^T v(t, \cdot) d\bar{\Pi}_t^2\right) \left(\int_0^T [u^2(t, \cdot) + v^2(t, \cdot)] dt\right)^{-1}$$

$$\hat{\varphi}_{2, T} = \left(\int_0^T u(t, \cdot) d\bar{\Pi}_t^2 - \int_0^T v(t, \cdot) d\bar{\Pi}_t^1\right) \left(\int_0^T [u^2(t, \cdot) + v^2(t, \cdot)] dt\right)^{-1}$$

Il est alors possible d'utiliser la technique de changement de temps : en effet :

$$\hat{\varphi}_{i, T} = \left(\int_0^T \langle a_i(t, \cdot), d\bar{\Pi}_t^\varphi \rangle\right) \left(\int_0^T \|a_i(t, \cdot)\|^2 dt\right)^{-1} + \varphi_i \quad ; \quad i=1, 2$$

où, $d\bar{\Pi}_t^\varphi = d\bar{\Pi}_t - \int_0^t [\varphi_1 a_1(s, \cdot) + \varphi_2 a_2(s, \cdot)] ds$, $t \geq 0$, et $(\bar{\Pi}_t^\varphi ; t \geq 0)$ est un mouvement brownien relativement à P_φ .

Ainsi, sous l'hypothèse que $P_\varphi[\int_0^{+\infty} (u^2(s, \cdot) + v^2(s, \cdot)) ds = +\infty] = 1$, posant :

$$\tau_H = \inf\{t \geq 0 : \int_0^t (u^2(s, \cdot) + v^2(s, \cdot)) ds \geq H\}$$

on a $\int_0^{\tau_H} \|a_i(t, \cdot)\|^2 dt = H$ et pour $i=1, 2$, $\int_0^{\tau_H} \langle a_i(t, \cdot), d\bar{\Pi}_t^\varphi \rangle$ est une variable gaussienne centrée de variance H . Par suite :

$$\tilde{\varphi}_{1, H} = \hat{\varphi}_{1, \tau_H} = -\frac{1}{H} \left[\int_0^{\tau_H} u(t, \cdot) d\bar{\Pi}_t^1 + \int_0^{\tau_H} v(t, \cdot) d\bar{\Pi}_t^2 \right]$$

$$\tilde{\varphi}_{2, H} = \hat{\varphi}_{2, \tau_H} = \frac{1}{H} \left[\int_0^{\tau_H} u(t, \cdot) d\bar{\Pi}_t^2 - \int_0^{\tau_H} v(t, \cdot) d\bar{\Pi}_t^1 \right]$$

sont des estimateurs sans biais de φ_1 et φ_2 respectivement, distribués suivant une loi gaussienne de variance $\frac{1}{H}$.

Nous pouvons sans difficulté généraliser ce résultat par la proposition suivante :

3.3.3. PROPOSITION : "Si a_1, a_2, \dots, a_d , $d \geq 2$, sont des applications telles que $\langle a_i(t, x), a_j(t, x) \rangle = 0$, $(t, x) \in \mathbb{R}_+ \times C$, $i \neq j$, et $P_{\varphi} [\int_0^{+\infty} \|a_i(t, \cdot)\|^2 dt = +\infty] = 1$ pour $i = \overline{1, d}$, les estimateurs de maximum de vraisemblance séquentiels

$$\hat{\varphi}_{i, H} = H^{-1} \int_0^{\tau_H^i} H \langle a_i(t, \cdot), d\Pi_t - a_0(t, \cdot) dt \rangle ; i = \overline{1, d}$$

avec $\tau_H^i = \inf\{t \geq 0 : \int_0^t \|a_i(s, \cdot)\|^2 ds \geq H\}$

de $\varphi_1, \dots, \varphi_d$ respectivement dans le modèle (28) où $\sigma(t, x) \equiv \frac{1}{p}$, sont des estimateurs sans biais distribués suivant une loi gaussienne de variance $1/H$."

3.3.4. Remarques.

a) Le modèle considéré dans la remarque 3.3.2.b) correspond au cas particulier de la proposition 3.3.3. :

$$p = d = 2 ; \|a_1(t, x)\|^2 = \|a_2(t, x)\|^2 ; a_0(t, x) = 0 .$$

b) Dans [16] F. DELBECQUE et J.P. QUADRAT ont étudié la situation où $p = 1$, $a_i(t, x) = a_i(t, x_t)$, $i = \overline{1, d}$, les fonctions a_i étant à supports disjoints dans $\mathbb{R}_+ \times \mathbb{R}$. Ils considèrent en particulier le cas où $a_i(t, x_t) = \chi_{A_i}(t, x_t)$ où $\{A_i ; i = \overline{1, d}\}$ forme une partition de $\mathbb{R}_+ \times \mathbb{R}$.

3.4. Estimation bayésienne des paramètres de dérive.

On considère le modèle (31) dans le cas scalaire i.e. $p = 1$. On suppose que la loi "a priori" de φ est une loi $N(m, \sigma^2)$. Le modèle

peut alors s'écrire :

$$(33) \quad \begin{cases} d\varphi_t = 0 & ; \quad t \geq 0 & ; \quad \varphi_0 = \varphi \sim N(m, \sigma^2) \\ dX_t = [a_0(t, X) + \varphi_t a_1(t, X)] dt + \sigma(t, X) dW_t & ; \quad t \geq 0 & ; \quad X_0 = 0 \end{cases}$$

C'est un modèle de filtrage non linéaire, le processus à estimer étant $\varphi_t \equiv \varphi$, le processus observé étant (X_t) . Engendrer dans ce modèle l'espérance conditionnelle de φ_t sachant $(X_s, 0 \leq s \leq t)$ revient exactement à déterminer une estimation bayésienne de φ .

A.N. SHIRYAYEV a étudié dans [47] un modèle de filtrage dont (33) est un cas particulier. Le théorème 1 de [47] permet d'énoncer la proposition suivante dont nous donnons toutefois une démonstration directe :

3.4.1. PROPOSITION : "La distribution conditionnelle de φ sachant $(X_t ; 0 \leq t \leq T)$ dans le modèle (33) est gaussienne de moyenne $\hat{\varphi}_T$ et de variance γ_T données par :

$$\hat{\varphi}_T = m \frac{\gamma_T}{\sigma^2} + \gamma_T \int_0^T \frac{a_1(s, \cdot)}{\sigma^2(s, \cdot)} (d\Pi_s - a_0(s, \cdot) ds)$$

et

$$\gamma_T = \frac{\sigma^2}{1 + \sigma^2 \int_0^T \frac{a_1^2(s, \cdot)}{\sigma^2(s, \cdot)} ds}$$

L'estimateur bayésien de φ dans (31) pour la loi "a priori" $N(m, \sigma^2)$ est :

$$\hat{\varphi}_T = [\sigma^{-2} + \int_0^T \frac{a_1^2(s, \cdot)}{\sigma^2(s, \cdot)} ds]^{-1} [\sigma^{-2} m + \int_0^T \frac{a_1(s, \cdot)}{\sigma^2(s, \cdot)} (d\Pi_s - a_0(s, \cdot) ds)] ."$$

Démonstration : Par hypothèse, la densité du couple $(\varphi, X_t ; 0 \leq t \leq T)$ par rapport à la mesure $\lambda \otimes \mu_*^T$, où λ est la mesure de Lebesgue sur \mathbb{R} , est donnée pour $u \in \mathbb{R}$ par :

$$q_T(u, \cdot) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(u-m)^2\right\} \times \left\{ \int_0^T \frac{a_0(t, \cdot) + u a_1(t, \cdot)}{\sigma^2(t, \cdot)} d\Pi_t - \frac{1}{2} \int_0^T \frac{(a_0(t, \cdot) + u a_1(t, \cdot))^2}{\sigma^2(t, \cdot)} dt \right\}$$

Alors la loi a postérieure de φ admet une densité par rapport à λ $q_T^*(\cdot)$ définie par :

$$q_T^*(u) = \frac{q_T(u, \cdot)}{\int_{\mathbb{R}} q_T(u, \cdot) du}$$

$q_T(u, \cdot)$ étant de la forme $K \exp Q_T(u, \cdot)$ où $Q_T(u, \cdot)$ est une forme quadratique en u , un calcul facile montre que :

$$\int_{\mathbb{R}} q_T(u, \cdot) du = \sigma^{-1} \cdot \gamma_T^{1/2} \exp \left\{ -\frac{m^2}{2\sigma^2} + \int_0^T \frac{a_0(t, \cdot)}{\sigma^2(t, \cdot)} d\Pi_t - \frac{1}{2} \int_0^T \frac{a_0^2(t, \cdot)}{\sigma^2(t, \cdot)} dt \right. \\ \left. + \frac{\gamma_T}{2} \left[\frac{m}{\sigma^2} + \int_0^T \frac{a_1(t, \cdot)}{\sigma^2(t, \cdot)} (d\Pi_t - a_0(t, \cdot) dt) \right]^2 \right\}.$$

On en déduit aisément que :

$$q_T^*(u) = \frac{1}{\sqrt{2\pi} \gamma_T^{1/2}} \exp \left\{ -\frac{1}{2\gamma_T} \left[u - \gamma_T (\sigma^{-2} m + \int_0^T \frac{a_1(t, \cdot)}{\sigma^2(t, \cdot)} (d\Pi_t - a_0(t, \cdot) dt)) \right]^2 \right\}$$

ce qui assure bien les conclusions de la proposition. ■

3.4.2. Remarque.

On voit que lorsque la variance σ^2 de la loi a priori tend vers l'infini, l'estimateur bayésien de φ converge vers l'estimateur :

$$\tilde{\varphi}_T = \left[\int_0^T \frac{a_1^2(s, \cdot)}{\sigma^2(s, \cdot)} ds \right]^{-1} \cdot \left[\int_0^T \frac{a_1(s, \cdot)}{\sigma^2(s, \cdot)} (d\Pi_s - a_0(s, \cdot) ds) \right]$$

qui n'est autre que l'estimateur de maximum de vraisemblance.

3.5. Quelques résultats numériques.

Nous présentons ici les résultats numériques d'expériences de simulation effectuées par A. MEGE-YTHIER et A. MUYSHONDT au cours d'un projet de 3e année ENSIMAG [38] dont le sujet leur avait été proposé

par J.L. SOLER et l'auteur.

La méthode a consisté à simuler des trajectoires d'une diffusion scalaire avec des paramètres donnés et à tenter ensuite de retrouver, par les estimations envisagées plus haut, ces paramètres. La simulation d'un mouvement brownien a été faite, selon une démarche suggérée par le théorème de DONSKER ([12] p.68), par la réalisation d'une marche aléatoire. On peut considérer que l'impureté du "bruit" obtenu nous a placés dans les conditions pratiques d'utilisation des techniques d'estimation envisagées, où le "bruit" n'est pas "blanc".

3.5.1. Estimation d'un paramètre de diffusion.

$$dX_t = \theta dW_t \quad ; \quad t \in [0, T] \quad ; \quad X_0 = 0$$

$$\hat{\sigma}_T^2 = \frac{1}{T} \sum_{i=1}^N \left(\Pi_{i \frac{T}{N}} - \Pi_{(i-1) \frac{T}{N}} \right)^2$$

$$E_{\theta}(\hat{\sigma}_T^2) = \theta^2 \quad ; \quad \sigma_{\theta}^2(\hat{\sigma}_T^2) = \frac{2\theta^4}{N}$$

Pour $\theta = 1$, $N = 10$ observations sur $[0, T] = [0, 1]$, 15 essais numériques ont donné les estimations suivantes :

$$\hat{\sigma}_T^2 = (1,356 ; 0,921 ; 0,851 ; 1,117 ; 0,702 ; 0,617 ; 1,405 ; 1,103 ; 0,834 ; 1,212 ; 1,503 ; 0,855 ; 0,532 ; 1,380 ; 1,380 ; 0,985)$$

	Théorique	Empirique
Moyenne	1	1,025
Variance	0,20	0,10

3.5.2. Estimation d'un paramètre de dérive.

$$a) \quad dX_t = \theta X_t dt + dW_t \quad ; \quad t \in [0, T] \quad ; \quad X_0 = 0$$

$$\hat{\theta}_T = \left(\int_0^T \Pi_t d\Pi_t \right) \left(\int_0^T \Pi_t^2 dt \right)^{-1} = \left(\frac{1}{2} (\Pi_T^2 - T) \right) \left(\int_0^T \Pi_t^2 dt \right)^{-1}$$

$$\hat{\theta}_{N,T} = \left(\sum_{i=1}^N \left(\Pi_{i \frac{T}{N}} - \Pi_{(i-1) \frac{T}{N}} \right) \Pi_{(i-1) \frac{T}{N}} \right) \left(\sum_{i=1}^N \frac{T}{N} \Pi_{(i-1) \frac{T}{N}}^2 \right)^{-1} .$$

Pour 100 observations par unité de temps, la moyenne des estimations sur 30 essais numériques, pour 3 valeurs différentes de θ , a été :

$\theta \backslash T$	1	2	3	5	8	11
-2	-3,62	-3,24	-2,65	-2,37	-2,21	-2,21
5	6,81	6,41	5,62	5,51	5,40	5,25
-20	-23,02	-22,20	-21,75	-21,62	-20,81	-20,21

La variance théorique asymptotique ($T \rightarrow +\infty$) de l'estimateur $\hat{\theta}_T$, lorsque θ est négatif, est $\frac{-2\theta}{T}$. On a comparé cette variance avec la variance empirique calculée pour les 30 essais numériques faits pour $\theta = -20$

T	2	3	5	8	11
Théorique	20	13,3	8	5	3,7
Empirique	33,5	23,2	13,4	11,2	5,3

$$\begin{aligned} \text{b) } dX_t &= \left(\sum_{i=0}^d \varphi_i t^i \right) X_t dt + dW_t \\ &= [\varphi_0(X_t) + \varphi_1(tX_t) + \dots + \varphi_d(t \cdot dX_t)] dt + dW_t ; \quad t \in [0, T] ; \\ X_0 &= 0 . \end{aligned}$$

$\hat{\varphi}_T$ est donné par la résolution de l'équation vectorielle :

$$A_T \hat{\varphi}_T = B_T$$

où $A_T = ((a_T^{ij}))_{i,j=1,d}$ avec $a_T^{ij} = \int_0^T t^{i+j-2} \Pi_t^2 dt$

et $B_T = ((b_T^i))_{i=1,d}$ avec $b_T^i = \int_0^T t^{i-1} \Pi_t d\Pi_t$

soit $b_T^i = \frac{1}{2} [T^{i-1} \Pi_T^2 - \frac{T^i}{i} - (i-1) \int_0^T t^{i-2} \Pi_t^2 dt]$.

Pour le cas où le polynôme en t intervenant dans le coefficient de dérive est de degré 1 (i.e. $d=1$) on a obtenu les résultats suivants :

$$* \varphi_0 = 10 \quad ; \quad \varphi_1 = -5$$

Moyenne des estimations sur 20 expériences :

$$\hat{\varphi}_0 = 12,3 \quad ; \quad \hat{\varphi}_1 = -5,91$$

$$* \varphi_0 = -3 \quad ; \quad \varphi_1 = 4$$

Moyenne des estimations sur 30 expériences :

$$\hat{\varphi}_0 = -2,7 \quad ; \quad \hat{\varphi}_1 = 3,89 .$$

Pour le cas où le polynôme en t est de degré 2 (i.e. $d=2$) les résultats ont été moins bons :

Moyenne des estimations sur 20 expériences :

$$* \varphi_0 = -5 \quad ; \quad \varphi_1 = 10 \quad ; \quad \varphi_2 = -3$$

$$\hat{\varphi}_0 = -4,57 \quad ; \quad \hat{\varphi}_1 = 2,06 \quad ; \quad \hat{\varphi}_2 = -1,88$$

$$* \varphi_0 = 1 \quad ; \quad \varphi_1 = -1 \quad ; \quad \varphi_2 = 2$$

$$\hat{\varphi}_0 = 3,4 \quad ; \quad \hat{\varphi}_1 = -2,5 \quad ; \quad \hat{\varphi}_2 = 3,2 .$$

3.5.3. Estimation simultanée d'un paramètre de dérive et d'un paramètre de diffusion.

$$dX_t = \varphi_1 X_t dt + \varphi_2 dW_t \quad ; \quad t \geq 0 \quad ; \quad X_0 = 0$$

$$\varphi_{2,T}^2 = \frac{1}{T} \sum_{i=1}^N \left(\Pi_{i \frac{T}{N}} - \Pi_{(i-1) \frac{T}{N}} \right)^2$$

$$\hat{\varphi}_{1,T} = \left(\sum_{i=1}^N \left(\Pi_{i \frac{T}{N}} - \Pi_{(i-1) \frac{T}{N}} \right) \Pi_{(i-1) \frac{T}{N}} \right) \left(\sum_{i=1}^N \frac{T}{N} \Pi_{(i-1) \frac{T}{N}}^2 \right)^{-1} .$$

Pour $\varphi_1 = -2$ et $\varphi_2 = 1$, on a trouvé pour 20 expériences, les moyennes empiriques $-2,30$ et $1,04$ des estimateurs $\hat{\varphi}_{1,10}$ et $\hat{\varphi}_{2,1}^2$, pour 100 observations par unité de temps.

Ces quelques résultats nous semblent encourageants et motiver à l'utilisation des techniques d'estimation que nous avons étudiées pour la modélisation de diffusions stochastiques.

Dans [16] F. DELBECQUE et J.P. QUADRAT présentent d'autres résultats numériques d'expériences de simulation et traitent l'identification d'une diffusion modélisant les apports d'eau à un réservoir hydraulique (E.D.F.) pour lequel ils disposaient d'observations semaine par semaine sur une vingtaine d'années.

CHAPITRE III

ESTIMATION DANS LES MODELES DIFFERENTIELS STOCHASTIQUES A COEFFICIENTS DEPENDANT NON LINEAIREMENT DES PARAMETRES INCONNUS

Un modèle classique de système dynamique stochastique en temps continu est celui défini par :

$$\begin{cases} dX_t = AX_t dt + Bu_t dt + CdW_t & ; t \geq 0 & ; X_0 = X(0) \\ dY_t = FX_t dt + Gd\xi_t & ; t \geq 0 & ; Y_0 = 0 \end{cases}$$

où A, B, C, F, G sont des matrices données de dimensions convenables, (X_t) représentant l'état d'un système à contrôler par un choix de (u_t) au vu d'une observation (Y_t) de (X_t) à travers un filtre bruité. Ce modèle a fait l'objet de nombreux travaux tant en ce qui concerne le problème de filtrage (voir par exemple R.S. BUCY et P.D. JOSEPH [14]), qu'en ce qui concerne le problème de contrôle et de filtrage simultanés (voir par exemple W.M. WONHAM [55]). Ici encore l'étude de ces problèmes suppose la connaissance complète des coefficients A, B, C, F, G des équations de la modélisation. Ce modèle, linéaire du point de vue des problèmes évoqués, cesse d'être linéaire quant au problème de l'estimation de paramètres inconnus au vu de l'observation de la seule sortie (Y_t) . En effet, \hat{X}_t désignant l'état filtré et \hat{W}_t étant l'innovation (qui est un mouvement brownien), Y_t admet la représentation

$$dY_t = F\hat{X}_t dt + Gd\hat{W}_t$$

où $F\hat{X}_t$ peut s'exprimer comme fonctionnelle de (Y_t) ; cette fonctionnelle, que nous déterminerons plus loin (§ 2.1), dépend de façon non linéaire des paramètres du modèle. Les résultats du chapitre précédent ne s'appliquent pas ici.

Dans la première partie de ce chapitre nous envisageons donc le problème d'estimation de paramètres inconnus apparaissant non linéairement dans les coefficients d'une équation de diffusion générale. Nous étudions encore la possibilité de découplage des estimations des paramètres de diffusion et des paramètres de dérive. Dans la deuxième partie du chapitre, en application de la première, nous considérons l'exemple du modèle de filtrage linéaire présenté plus haut. Nous terminons ce chapitre par quelques remarques sur les problèmes de modélisation de systèmes et d'identification de paramètres et quelques suggestions de prolongements possibles à notre travail.

§ 1. - ESTIMATION DANS UN MODELE DIFFERENTIEL STOCHASTIQUE A COEFFICIENTS DEPENDANT DIFFERENTIABLEMENT DES PARAMETRES INCONNUS.

On s'intéresse ici au modèle présenté sous sa forme la plus générale dans le chapitre I, § 1.4 : Φ étant un sous-ensemble de \mathbb{R}^d (d entier donné, $d \geq 1$), $\{(a_\varphi, \sigma_\varphi) ; \varphi \in \Phi\}$ une famille de couples d'applications telles que, pour tout $\varphi \in \Phi$ les hypothèses (D1) et (D2) sont satisfaites, la structure statistique du problème d'estimation de $\varphi = (\varphi_1, \dots, \varphi_d)$ pour une observation en temps continu sur $[0, T]$, est définie par :

$$(C, S_T, \{P_\varphi^T ; \varphi \in \Phi\})$$

où, pour tout $\varphi \in \Phi$, $P_\varphi^T = P_{\varphi/S_T}$, avec :

$$(1) \quad \frac{dP_\varphi}{dQ_\varphi} / S_T = \exp\left\{\int_0^T \langle \sigma_\varphi^{-1}(s, \cdot) a_\varphi(s, \cdot), \sigma_\varphi^{-1}(s, \cdot) d\Pi_s \rangle - \frac{1}{2} \int_0^T \|\sigma_\varphi^{-1}(s, \cdot) a_\varphi(s, \cdot)\|^2 ds\right\}$$

Q_φ étant la solution faible de l'équation :

$$\begin{cases} dX_t = \sigma_\varphi(t, X) dW_t & ; \quad t \geq 0 \\ X_0 = 0 \end{cases}$$

Supposons que pour un r entier, $1 \leq r \leq d$, $\varphi_\Delta = (\varphi_1, \dots, \varphi_r)$ soit le vecteur des paramètres dont dépend σ_φ i.e. $\sigma_\varphi = \sigma_{\varphi'} = \tilde{\sigma}_{\varphi_\Delta}$ si

$\varphi_{\Delta} = \varphi'_{\Delta}$, $(\varphi, \varphi') \in \Phi^2$. Nous montrons, sous certaines hypothèses, qu'il est possible de découper l'estimation de φ_{Δ} de celle de $\varphi_{\Delta}^{-} = (\varphi_{r+1}, \dots, \varphi_d)$ en ce sens que φ_{Δ} est fortement identifiable et que, φ_{Δ} étant supposé connu (parce qu'identifié sur un intervalle de temps fini), φ_{Δ}^{-} est asymptotiquement identifiable par maximum de vraisemblance par une famille d'estimateurs asymptotiquement normalement distribuée.

L'utilisation pratique des résultats est suggérée par les énoncés. En effet, qu'il s'agisse des paramètres de diffusion ou des paramètres de dérive, ils montrent que l'identification repose sur la mise en oeuvre de méthodes numériques pour résoudre une équation vectorielle, dont les coefficients dépendent de la trajectoire observée. Nous indiquons un algorithme permettant d'envisager la résolution effective des équations mise en jeu. Les résultats de ce paragraphe ont été présentés dans [34] A. LE BRETON.

1.1. Estimation des paramètres du coefficient de diffusion.

Rappelons le résultat du lemme II, 3.1.1. : pour tout $\varphi \in \Phi$ et tout $t \in]0, T]$ on a :

$$(2) \quad P_{\varphi}^T - \lim_{N \rightarrow +\infty} \hat{\sigma}_t^N = \int_0^t \sigma_{\varphi_{\Delta}}(s, \cdot) \cdot {}^t \sigma_{\varphi_{\Delta}}(s, \cdot) ds$$

où $\hat{\sigma}_t^N(x) = \sum_{i=1}^{2N} (x_{\frac{i}{2N}t} - x_{\frac{(i-1)}{2N}t})^2 \otimes$ pour $x = (x_t; t \geq 0) \in C$, $N \geq 1$.

On est alors conduit à envisager l'identification de la valeur φ_{Δ}^* par la résolution d'un système d'équations en φ_{Δ} du type :

$$(3) \quad \begin{cases} \int_0^{t_i} [\sigma_{\varphi_{\Delta}}(s, x) \cdot {}^t \sigma_{\varphi_{\Delta}}(s, x)]_{k_j^i, \ell_j^i} ds = [\hat{\sigma}_{t_i}^N]_{k_j^i, \ell_j^i} ; x \in C ; N \geq 1 \\ i \in \{1, 2, \dots, k\} , j \in \{1, \dots, j(i)\} \end{cases}$$

où $0 < t_1 < t_2 < \dots < t_k \leq T$,

$$1 \leq k_j^i \leq \ell_j^i \leq p ; (k_j^i, \ell_j^i) \neq (k_{j'}^i, \ell_{j'}^i) \text{ si } j \neq j' , (j, j') \in \{1, \dots, j(i)\}^2$$

et k et $j(i)$, $i \in \{1, \dots, k\}$ sont choisis de telle manière que $\sum_{i=1}^k j(i) = r$.

Nous écrirons, pour simplifier les notations, le système (3) sous la forme de l'équation vectorielle dans \mathbb{R}^r :

$$(4) \quad F(\varphi_\Delta, x) = \gamma^N(x) \quad ; \quad x \in C \quad ; \quad N \geq 1$$

où, si (θ_1, θ_2) est une paramétrisation (injective) de $\{(t_i, (k_j^i, \ell_j^i)) ; i = \overline{1, k}, j \in \{1, \dots, j(i)\}\}$ sur $\{1, \dots, r\}$ et, pour $u = \overline{1, r}$,

$$F_u(\varphi_\Delta, x) = \left(\int_0^{\theta_1(u)} \sigma_{\varphi_\Delta}(s, x) \cdot {}^t \sigma_{\varphi_\Delta}(s, x) ds \right)_{\theta_2(u)} \quad (\text{resp. } \gamma_u^N(x) = (\hat{\sigma}_{\theta_1(u)}^N)_{\theta_2(u)}) ,$$

on a posé :

$$F(\varphi_\Delta, x) = {}^t(F_1(\varphi_\Delta, x), \dots, F_r(\varphi_\Delta, x)) \quad (\text{resp. } \gamma^N(x) = {}^t(\gamma_1^N(x), \dots, \gamma_r^N(x)) .$$

Compte-tenu de (2), pour tout $\varphi = (\varphi_\Delta, \varphi_{\bar{\Delta}}) \in \Phi$ tel que $\varphi_\Delta = \varphi_\Delta^*$, on a :

$$(5) \quad P_\varphi^T \text{-} \lim_{N \rightarrow +\infty} \gamma^N(\cdot) = F(\varphi_\Delta^*, \cdot) .$$

On peut alors s'attendre à ce que, sous des hypothèses convenables sur l'application F , la valeur φ_Δ^* soit fortement identifiable en probabilité sur l'intervalle de temps $[0, T]$, au sens de la définition I, 1.4.1 et, plus particulièrement, qu'il existe une suite $\{\hat{\varphi}_{\Delta, T}^N ; N \geq 1\}$ de statistiques définies sur la structure $(C, S_T, \{P_\varphi^T ; \varphi \in \Phi\})$ telle que, pour tout $\varphi = (\varphi_\Delta, \varphi_{\bar{\Delta}}) \in \Phi$ où $\varphi_\Delta = \varphi_\Delta^*$, on ait :

$$(6) \quad \begin{cases} P_\varphi^T \text{-} \lim_{N \rightarrow +\infty} \hat{\varphi}_{\Delta, T}^N = \varphi_\Delta^* \\ \lim_{N \rightarrow +\infty} P_\varphi^T [F(\hat{\varphi}_{\Delta, T}^N, \cdot) = \gamma^N(\cdot)] = 1 \end{cases}$$

1.1.1. PROPOSITION : "On suppose qu'il existe un ensemble d'instants $0 < t_1 < \dots < t_k \leq T$ et des ensembles d'indices (k_j^i, ℓ_j^i) , $i = \overline{1, k}$, pour lesquels il existe des fonctions aléatoires $(\varphi_\Delta, x) \rightsquigarrow \tilde{F}(\varphi_\Delta, x)$ et $(\varphi_\Delta, x) \rightsquigarrow \tilde{F}^{(1)}(\varphi_\Delta, x) = (\tilde{F}_i^{(1)}(\varphi_\Delta, x) ; i = \overline{1, r})$ à trajectoires continues telles que :

- (i) $\tilde{F}(\varphi_\Delta, \cdot) = F(\varphi_\Delta, \cdot)$ presque sûrement, pour tout φ_Δ ,
- (ii) $\frac{\partial}{\partial \varphi_i} \tilde{F}(\varphi_\Delta, x) = \tilde{F}_i^{(1)}(\varphi_\Delta, x)$ pour tout (x, φ_Δ) , $\tilde{F}_i^{(1)}(\varphi_\Delta^*, x)$ étant non singulière,

(iii) il existe une famille $\{\epsilon(\eta) ; \eta > 0\}$ de variables aléatoires positives finies, S_T -mesurables, convergeant en probabilité vers 0 lorsque η tend vers 0, telle que :

$$\sup_{\|\varphi_{\Delta} - \varphi_{\Delta}^*\| \leq \eta} \|\tilde{F}^{(1)}(\varphi_{\Delta}, \cdot) - \tilde{F}^{(1)}(\varphi_{\Delta}^*, \cdot)\| \leq \epsilon(\eta) .$$

Alors φ_{Δ}^* est fortement identifiable en probabilité sur l'intervalle $[0, T]$ par une suite de statistiques $\{\hat{\varphi}_{\Delta, T}^N ; N \geq 1\}$ satisfaisant aux conditions (6). En outre, si $\{\tilde{\varphi}_{\Delta, T}^N ; N \geq 1\}$ est une autre suite de statistiques vérifiant (6) on a :

$$\lim_{N \rightarrow +\infty} P_{\varphi}^T[\tilde{\varphi}_{\Delta, T}^N = \hat{\varphi}_{\Delta, T}^N] = 1$$

pour tout $\varphi = (\varphi_{\Delta}, \varphi_{-}) \in \Phi$ tel que $\varphi_{\Delta} = \varphi_{\Delta}^*$."

Démonstration : On note F et $F^{(1)}$ pour \tilde{F} et $\tilde{F}^{(1)}$ respectivement. Pour $N \in \mathbb{N}^*$ et $K \in \mathbb{N}^*$ on considère l'événement S_T -mesurable :

$$\Omega_N(K) = \{x \in C : \|[F^{(1)}(\varphi_{\Delta}^*, x)]^{-1}\| \|F(\varphi_{\Delta}^*, x) - \gamma^N(x)\| \leq \frac{1}{3K} , \|[F^{(1)}(\varphi_{\Delta}^*, x)]^{-1}\| \cdot \epsilon(\frac{1}{K}, x) \leq \frac{1}{3}\} .$$

On a, pour $\varphi = (\varphi_{\Delta}^*, \varphi_{-})$,

$$P_{\varphi}[\bigcap_N \Omega_N(K)] \leq P_{\varphi}[\|[F^{(1)}(\varphi_{\Delta}^*, \cdot)]^{-1}\| \|F(\varphi_{\Delta}^*, \cdot) - \gamma^N(\cdot)\| > \frac{1}{3K}] + P_{\varphi}[\|[F^{(1)}(\varphi_{\Delta}^*, \cdot)]^{-1}\| \cdot \epsilon(\frac{1}{K}) > \frac{1}{3}]$$

où, compte-tenu de (5) on a pour tout $K \in \mathbb{N}^*$:

$$\lim_{N \rightarrow +\infty} P_{\varphi}[\|[F^{(1)}(\varphi_{\Delta}^*, \cdot)]^{-1}\| \|F(\varphi_{\Delta}^*, \cdot) - \gamma^N(\cdot)\| > \frac{1}{3K}] = 0 .$$

Alors on a :

$$\limsup_{N \rightarrow +\infty} P_{\varphi}[\bigcap_N \Omega_N(K)] \leq P_{\varphi}[\|[F^{(1)}(\varphi_{\Delta}^*, \cdot)]^{-1}\| \cdot \epsilon(\frac{1}{K}) > \frac{1}{3}] ,$$

puis, compte-tenu de (iii),

$$\lim_{K \rightarrow +\infty} \limsup_{N \rightarrow +\infty} P_{\varphi}[\bigcap_N \Omega_N(K)] = 0$$

ou

$$\lim_{K \rightarrow +\infty} \liminf_{N \rightarrow +\infty} P_{\varphi}[\Omega_N(K)] = 1 .$$

Soit, pour $N \in \mathbb{N}^*$ et $K \in \mathbb{N}^*$ fixés, $x \in \Omega_N(K)$: on définit l'application $\psi_N(x, \cdot)$ sur $\{u \in \mathbb{R}^r : \|u\| \leq \frac{1}{K}\}$ par :

$$\psi_N(x, u) = u - [F^{(1)}(x, \varphi_\Delta^*)]^{-1}(F(x, \varphi_\Delta^* + u) - \gamma^N(x)) .$$

On montre sans difficulté que $\psi_N(x, \cdot)$ est une application contractante à valeurs dans $\{u \in \mathbb{R}^r : \|u\| \leq \frac{1}{K}\}$. Alors $\psi_N(x, \cdot)$ admet un point fixe unique i.e. il existe $\bar{u}_N(x)$ unique tel que $\|\bar{u}_N(x)\| \leq \frac{1}{K}$ et

$$0 = [F^{(1)}(x, \varphi_\Delta^*)]^{-1}(F(x, \varphi_\Delta^* + \bar{u}_N(x)) - \gamma^N(x))$$

puis $F(x, \varphi_\Delta^* + \bar{u}_N(x)) = \gamma^N(x)$.

Posons $\Omega_N = \bigcup_{K \in \mathbb{N}^*} \Omega_N(K) \in S_T$: d'après ce qui précède, pour tout $x \in \Omega_N$ il existe un entier $K(x) \geq 1$ tel que l'équation $F(x, \varphi_\Delta) = \gamma^N(x)$ admette une solution unique $\tilde{\varphi}_\Delta^N(x)$ dans la boule $\|\varphi_\Delta - \varphi_\Delta^*\| \leq \frac{1}{K}$.

Nous sommes donc dans les conditions d'application du lemme I, 2.1.2 en y faisant :

$$(\Omega, \mathcal{A}) = (\Omega_N, S_T \cap \Omega_N)$$

$$\ell(x, \varphi_\Delta) = F(x, \varphi_\Delta) - \gamma^N(x)$$

et $h(x) \equiv \varphi_\Delta^*$.

Par suite, l'application $x \rightsquigarrow \tilde{\varphi}_\Delta^N(x)$ définie sur $(\Omega_N, S_T \cap \Omega_N)$ est mesurable et posant :

$$\hat{\varphi}_\Delta^N(x) = \begin{cases} \tilde{\varphi}_\Delta^N(x) & \text{si } x \in \Omega_N \\ \varphi_\Delta^\circ & \text{si } x \notin \Omega_N \end{cases}$$

où φ_Δ° est une valeur quelconque du paramètre, l'application $\hat{\varphi}_\Delta^N$ est S_T -mesurable.

On démontre alors, suivant la même démarche que dans I, 2.3.3, que la suite $\{\hat{\varphi}_\Delta^N ; N \geq 1\}$ vérifie les conditions (6) et est "unique avec probabilité tendant vers 1". ■

1.1.2. Remarques :

a) Les hypothèses (i) et (ii) correspondent bien sûr à des hypothèses de différentiabilité du coefficient de diffusion σ_φ par rapport à φ_Δ

φ_{Δ} , le problème de l'existence d'une "version régulière" de $(\varphi_{\Delta}, x) \rightsquigarrow F(\varphi_{\Delta}, x)$ étant de même nature que celui de l'existence d'une "version régulière" d'une fonction de vraisemblance que nous envisageons plus loin à la lumière du I. § 3.1.

b) Dans le cas où F prend la forme :

$$F(\varphi_{\Delta}, x) = \sum_{i=1}^r \varphi_i F_i(x)$$

ce qui peut, par exemple, provenir d'une paramétrisation de $\sigma_{\varphi_{\Delta}}$ du type

$$\sigma_{\varphi_{\Delta}}(t, x) = \sqrt{\varphi_j} \cdot \sigma_{i,j}(t, x) \quad ; \quad i, j = \overline{1, p} ,$$

les hypothèses se traduisent par :

$$\det(F_1(x), \dots, F_r(x)) \neq 0 .$$

c) Lorsque F ne dépend pas de x , par exemple parce que $\sigma_{\varphi_{\Delta}}$ n'en dépend pas, les hypothèses (i) et (ii) consistent à supposer que $F(\cdot)$ est un homéomorphisme local en φ_{Δ}^* . Dans le cas particulier où $\sigma_{\varphi_{\Delta}}$ ne dépend ni de x ni de t on a bien sûr, compte-tenu du lemme II, I.2.2., que $\sigma_{\varphi_{\Delta}} \cdot \sigma_{\varphi_{\Delta}}^t$ est fortement identifiable presque sûrement sur tout intervalle de temps fini $[0, t]$, par la suite de statistiques $\{\frac{1}{t} \hat{\sigma}_t^N ; N \geq 1\}$.

d) Dans [16] F. DELBECQUE et J.P. QUADRAT ont proposé, pour le cas où :

$$\sigma_{\varphi_{\Delta}}(t, x) = \sigma_{\varphi_{\Delta}}(t, x_t) = \varphi_{\Delta} \in \mathbb{R} \quad , \quad (t, x_t) \in \mathbb{R}_+ \times \mathbb{R} ,$$

une estimation sans biais de φ_{Δ}^2 faisant intervenir des temps d'arrêt du processus observé, dans l'esprit du II. § 3.3.

1.2. Estimation des paramètres du coefficient de dérive.

Nous supposons désormais que le coefficient de diffusion σ ne dépend pas de paramètres inconnus, éventuellement parce que le paramètre φ_{Δ} y figurant aurait été identifié sur un intervalle de temps fini.

Le problème est alors d'estimer le paramètre vectoriel φ_{Δ} constitué des composantes de φ apparaissant dans le seul coefficient de dérive. Afin de simplifier les notations nous noterons à nouveau $\varphi_{\Delta} = \varphi \in \Phi \subset \mathbb{R}^d$, la structure statistique du problème d'estimation :

$$(C, S_T, \{P_{\varphi}^T; \varphi \in \Phi\})$$

étant ici dominée par la mesure $Q^T = Q/S_T$, où Q est la solution faible de l'équation :

$$\begin{cases} dX_t = \sigma(t, X) dW_t & ; \quad t \geq 0 \\ X_0 = 0 & , \end{cases}$$

avec

$$(7) \quad L_T(\varphi, \cdot) = \text{Log} \frac{dP_{\varphi}}{dQ} / S_T(\cdot) = \int_0^T \langle \sigma^{-1}(s, \cdot) a_{\varphi}(s, \cdot), \sigma^{-1}(s, \cdot) d\Pi_s \rangle - \frac{1}{2} \int_0^T \|\sigma^{-1}(s, \cdot) a_{\varphi}(s, \cdot)\|^2 ds$$

l'intégrale stochastique étant prise relativement à $(C, (S_t)_{t \geq 0}, Q)$.

On fait les hypothèses suivantes, φ^* désignant la vraie valeur de φ :

(K1) Pour tout $(t, x) \in \mathbb{R}_+ \times C$, l'application $\varphi \rightsquigarrow a_{\varphi}(t, x)$ est trois fois différentiable de dérivées successives $a_{\varphi}^{(k)}(t, x)$, $k = \overline{1, 3}$, et il existe un processus $(U_t)_{t \geq 0}$ basé sur C tel que, r désignant un réel $r \geq 2$, $r > d$,
 $\int_0^T E_Q(|U_t|^r) dt < +\infty$, $T > 0$
 et $\sup_{\varphi \in \Phi} \|\sigma^{-1}(t, x) a_{\varphi}^{(k)}(t, x)\|^2 \leq U_t(x)$; $(t, x) \in \mathbb{R}_+ \times C$, $k = \overline{1, 3}$.

(K2) $\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \sigma^{-1}(t, \cdot) a_{\varphi^*}^{(1)}(t, \cdot) \cdot a_{\varphi^*}^{(1)}(t, \cdot) \cdot \sigma^{-1}(t, \cdot) dt = \Gamma$,

P_{φ^*} presque sûrement, et

$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T E_{\varphi^*}(\sigma^{-1}(t, \cdot) a_{\varphi^*}^{(1)}(t, \cdot) \cdot a_{\varphi^*}^{(1)}(t, \cdot) \cdot \sigma^{-1}(t, \cdot)) dt = \Gamma$

où Γ est une matrice symétrique définie positive.

(K3) $\lim_{T \rightarrow +\infty} \frac{1}{T^2} \int_0^T E_{\varphi^*} \|\sigma^{-1}(t, \cdot) a_{\varphi^*}^{(2)}(t, \cdot)\|^2 dt = 0$.

Nous assurons d'abord l'existence d'une version régulière de la fonction de Log-vraisemblance $L_T(\varphi, \cdot)$: on définit les fonctions aléatoires $L_T^{(1)}(\varphi, \cdot)$ et $L_T^{(2)}(\varphi, \cdot)$, basées sur $\mathfrak{F} \times C$, par :

$$(8) \quad L_T^{(1)}(\varphi, \cdot) = \left(\int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) (d\Pi_s - a_\varphi(s) ds) \rangle ; i=1, \overline{d} \right)$$

et

$$(9) \quad L_T^{(2)}(\varphi, \cdot) = \left(\int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) (d\Pi_s - a_\varphi(s) ds) \right. \\ \left. - \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_j} a_\varphi(s, \cdot) \rangle ; i, j=1, \overline{d} \right) .$$

Remarquons que la fonction aléatoire $(Y_t ; t \geq 0)$ où :

$$Y_t = \int_0^t \sigma^{-1}(s, \cdot) d\Pi_s$$

définit un processus de mouvement brownien relativement à la mesure Q .

Comme elle admet la représentation :

$$Y_t = \int_0^t \sigma^{-1}(s, \cdot) (d\Pi_s - a_\varphi^*(s, \cdot) ds) + \int_0^t \sigma^{-1}(s, \cdot) a_\varphi^*(s, \cdot) ds ; t \geq 0 ,$$

où, relativement à la mesure P_{φ^*} , le processus $(\Pi_t^* ; t \geq 0)$ où :

$$\Pi_t^* = \int_0^t \sigma^{-1}(s, \cdot) (d\Pi_s - a_\varphi^*(s, \cdot) ds) ; t \geq 0 ,$$

est un processus de mouvement brownien, d'après J. MEMIN [39], 1ère partie, lemme 2.6., compte-tenu de l'équivalence des mesures Q et

P_{φ^*} on peut aussi bien définir les expressions (7), (8) et (9) respectivement par :

$$(7') \quad L_T(\varphi, \cdot) = \int_0^T \langle \sigma^{-1}(s, \cdot) a_\varphi(s, \cdot), d\Pi_s^* \rangle - \frac{1}{2} \int_0^T \|\sigma^{-1}(s, \cdot) (a_\varphi(s, \cdot) - a_{\varphi^*}(s, \cdot))\|^2 ds \\ + \frac{1}{2} \int_0^T \|\sigma^{-1}(s, \cdot) a_{\varphi^*}(s, \cdot)\|^2 ds$$

$$(8') \quad L_{T,i}^{(1)}(\varphi, \cdot) = \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_\varphi(s, \cdot), d\Pi_s^* \rangle \\ + \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) (a_{\varphi^*}(s, \cdot) - a_\varphi(s, \cdot)) \rangle ds$$

$$(9') \quad L_{T,i,j}^{(2)}(\varphi, \cdot) = \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} a_\varphi(s, \cdot), d\Pi_s^* \rangle \\ - \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_j} a_\varphi(s, \cdot) \rangle ds \\ + \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) (a_{\varphi^*}(s, \cdot) - a_\varphi(s, \cdot)) \rangle ds$$

les intégrales stochastiques étant prises relativement à P_{φ^*} .

On démontre maintenant un lemme :

1.2.1. LEMME : "Sous l'hypothèse (K1) il existe des fonctions aléatoires $\varphi \rightsquigarrow \hat{L}_T(\varphi, .)$, $\varphi \rightsquigarrow \hat{L}_T^{(1)}(\varphi, .) = (\hat{L}_{T,i}^{(1)}(\varphi, .) ; i=\overline{1,d})$ et $\varphi \rightsquigarrow \hat{L}_T^{(2)}(\varphi, .) = (\hat{L}_{T,i,j}^{(2)}(\varphi, .) ; i,j=\overline{1,d})$, à trajectoires continues telles que l'on ait :

(i) pour tout $\varphi \in \Phi$, Q-presque sûrement :

$$\hat{L}_T(\varphi, .) = L_T(\varphi, .) \quad \text{et} \quad \hat{L}_T^{(k)}(\varphi, .) = L_T^{(k)}(\varphi, .) , \quad k=\overline{1,2}$$

(ii) pour $i,j = \overline{1,d}$ et tout $x \in C$:

$$\hat{L}_{T,i}^{(1)}(\varphi, x) = \frac{\partial}{\partial \varphi_i} \hat{L}_T(\varphi, x) \quad \text{et} \quad \hat{L}_{T,i,j}^{(2)}(\varphi, x) = \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} \hat{L}_T(\varphi, x) ."$$

Démonstration : Elle consiste à vérifier que nous sommes dans les conditions d'application de la proposition I, 3.1.1. L'hypothèse (K1) permet de vérifier sans difficulté, en utilisant en particulier le théorème de la convergence dominée, que les fonctions aléatoires figurant dans les intégrales stochastiques et ordinaires des expressions (7), (8) et (9) satisfont aux conditions correspondantes des hypothèses (G1) et (G2) de cette proposition relatives aux processus $(\Psi_\varphi(t) ; t \geq 0)$ et $(\Lambda_\varphi(t) ; t \geq 0)$. Alors $(Y_t ; t \geq 0)$ étant un mouvement brownien relativement à Q , il existe des fonctions aléatoires $\hat{L}_T^{(1)}(\varphi, .)$ et $\hat{L}_T^{(2)}(\varphi, .)$ à trajectoires continues telles que l'on ait :

$$\hat{L}_T^{(k)}(\varphi, .) = L_T^{(k)}(\varphi, .) , \quad k=\overline{1,2} , \quad Q\text{-presque sûrement}$$

et
$$\hat{L}_{T,i,j}^{(2)}(\varphi, x) = \frac{\partial}{\partial \varphi_i} \hat{L}_{T,j}^{(1)}(\varphi, x) , \quad (\varphi, x) \in \Phi \times C , \quad i,j=\overline{1,d} .$$

Procédant comme dans la démonstration de la proposition I, 3.1.1. on définit alors $\hat{L}_T(\varphi, .)$ par intégration de $\hat{L}_T^{(1)}(\varphi, .)$ et on montre de même que $\hat{L}_T(\varphi, .)$ est une version de $L_T(\varphi, .)$ défini par (7). ■

Dans la suite nous considérerons toujours une version régulière $\hat{L}_T(\varphi, .)$ de la fonction de Log-vraisemblance et pour alléger les notations

nous écrirons L_T , $L_T^{(1)}$ et $L_T^{(2)}$ au lieu de \hat{L}_T , $\hat{L}_T^{(1)}$ et $\hat{L}_T^{(2)}$ respectivement.

On énonce maintenant une proposition assurant l'identifiabilité de φ^* :

1.2.2. PROPOSITION : "Sous les hypothèses (K1), (K2), (K3) et (E4), la valeur φ^* est asymptotiquement identifiable en probabilité. De manière précise, il existe une famille $\{\hat{\varphi}_T ; T > 0\}$ d'estimateurs de φ vérifiant :

- (i) pour tout $T > 0$, $\hat{\varphi}_T$ est S_T -mesurable,
- (ii) $\lim_{T \rightarrow +\infty} P_{\varphi^*} [L_T^{(1)}(\hat{\varphi}_T, \cdot) = 0] = 1$,
- (iii) $P_{\varphi^*} - \lim_{T \rightarrow +\infty} \hat{\varphi}_T = \varphi^*$.

En outre, si $\{\tilde{\varphi}_T ; T > 0\}$ est une autre famille d'estimateurs vérifiant ces 3 conditions, alors :

- (iv) $\lim_{T \rightarrow +\infty} P_{\varphi^*} [\tilde{\varphi}_T = \hat{\varphi}_T] = 1$.

Enfin, si $\{\hat{\varphi}_T ; T > 0\}$ est une famille d'estimateurs vérifiant (i), (ii) et (iii) on a :

- (v) $\mathcal{L}_{P_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{1/2}(\hat{\varphi}_T - \varphi^*) = N(0, \Gamma^{-1})$, Γ étant définie par (K2),
- (vi) $\mathcal{L}_{P_{\varphi^*}} - \lim_{T \rightarrow +\infty} 2[L_T(\hat{\varphi}_T, \cdot) - L_T(\varphi^*, \cdot)] = \chi_d^2$.

Démonstration : Elle consiste évidemment à vérifier que les hypothèses (E1), (E2) et (E3) de la proposition I, 2.2.1. sont satisfaites.

On vérifie d'abord (E1) : on a, d'après (8')

$$L_{T,i}^{(1)}(\varphi^*, \cdot) = \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_{\varphi^*}(s, \cdot), d\Pi_s^* \rangle, \quad i = \overline{1, d}$$

où $(\Pi_t^* ; t \geq 0)$ est un mouvement brownien relativement à P_{φ^*} .

Alors, comme :

$$E_{\varphi^*} [|L_{T,i}^{(1)}(\varphi^*, \cdot)|^2] \leq p E_{\varphi^*} \int_0^T \| \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_{\varphi^*}(s, \cdot) \|^2 ds$$

et que, d'après (K2), on a :

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T E_{\varphi^*} \| \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_{\varphi^*}(s, \cdot) \|^2 ds = \Gamma_{ii} .$$

On a aussi $\lim_{T \rightarrow +\infty} E_{\varphi^*} [| \frac{1}{T} L_{T,i}^{(1)}(\varphi^*, \cdot) |^2] = 0$, ce qui assure que (E1) est vérifiée.

On s'intéresse maintenant à l'hypothèse (E2) : de même que précédemment on montre, à l'aide de (K3), que :

$$P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} a_{\varphi^*}(s, \cdot) , d\Pi_s^* \rangle = 0 .$$

D'après (9'), pour montrer (E2) il suffit donc de vérifier que :

$$P_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_{\varphi^*}(s, \cdot) , \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_j} a_{\varphi^*}(s, \cdot) \rangle ds = \Gamma_{ij} ,$$

ce qui est acquis en vertu de (K2).

L'hypothèse (E3) est satisfaite compte-tenu de l'hypothèse (K2) et du théorème I, 3.2.1.

Les conclusions de la proposition résultent alors de la proposition I, 2.2.1. ■

1.2.3. Remarque :

L'hypothèse (E4) de la proposition I, 2.2.1. figure, en tant que telle, comme hypothèse dans l'énoncé précédent alors que nous donnons des conditions suffisantes (K1, K2, K3) pour assurer les hypothèses (E1, E2, E3) de la dite proposition. Nous n'avons pas su dégager une condition générale qui assure que (E4) est ici vérifiée, ayant pensé un temps que les hypothèses (K1, K2, K3) y suffisaient, ce que nous ne sommes pas parvenus à démontrer.

Nous nous proposons toutefois de cerner le problème et de montrer qu'il se résoud dans un cas particulier : l'expression mise en jeu s'écrit pour $i, j = \overline{1, d}$,

$$\begin{aligned} \frac{1}{T} [L_{T,i,j}^{(2)}(\varphi, \cdot) - L_{T,i,j}^{(2)}(\varphi^*, \cdot)] &= \frac{1}{T} \int_0^T \langle \sigma^{-1}(s, \cdot) \left(\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} a_\varphi(s, \cdot) - \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} a_{\varphi^*}(s, \cdot) \right), d\Pi_s^* \rangle \\ &+ \frac{1}{T} \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) (a_{\varphi^*}(s, \cdot) - a_\varphi(s, \cdot)) \rangle ds \\ &- \frac{1}{T} \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_\varphi(s, \cdot), \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_j} a_\varphi(s, \cdot) \rangle ds \\ &+ \frac{1}{T} \int_0^T \langle \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_i} a_{\varphi^*}(s, \cdot), \sigma^{-1}(s, \cdot) \frac{\partial}{\partial \varphi_j} a_{\varphi^*}(s, \cdot) \rangle ds . \end{aligned}$$

A l'aide des hypothèses K 1 et K 2 on montre aisément que la somme $\sum_{i,j}^\varphi(T, \cdot)$ des 3 derniers termes du second membre vérifie :

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} |\sum_{i,j}^\varphi(T, \cdot)| \leq \eta M_T^{i,j} \quad \text{où } M_T^{i,j} \text{ est une variable aléatoire d'espérance bornée indépendamment de } T .$$

Si l'intégrale stochastique du second membre vérifiait la même propriété, l'hypothèse (E4) serait vérifiée. Le problème est donc celui de la majoration d'une intégrale stochastique :

$$\frac{1}{T} \int_0^T [V_\varphi(s) - V_{\varphi^*}(s)] dW_s$$

où $\{W_t; t \geq 0\}$ est un mouvement brownien,

* $\{V_\varphi(t); t \geq 0\}; \varphi \in \Phi$ est une famille de processus.

Plaçons-nous dans le cas scalaire (i.e. $V_\varphi(t) \in \mathbb{R}, W_t \in \mathbb{R}, \varphi \in \mathbb{R}$) où pour tout $\varphi \in \Phi$, le processus $\{V_\varphi(t); t \geq 0\}$ est presque sûrement à trajectoires dérivables, de dérivée $\{\tilde{V}_\varphi(t); t \geq 0\}$ et supposons que les applications $\varphi \rightsquigarrow V_\varphi(t)$ et $\varphi \rightsquigarrow \tilde{V}_\varphi(t)$ sont dérivables de dérivées respectives $V_\varphi^{(1)}(t)$ et $\tilde{V}_\varphi^{(1)}$ telles que :

$$\sup_{\varphi \in \Phi} \|V_\varphi^{(1)}(t)\| \leq U(t)$$

et

$$\sup_{\varphi \in \Phi} \|\tilde{V}_\varphi^{(1)}(t)\| \leq \tilde{U}(t)$$

où $\frac{1}{T} E(U_T^2) \leq K < +\infty$ et $\frac{1}{T} E[(\int_0^T \tilde{U}_t dt)^2] \leq K$.

On a alors, en intégrant par parties :

$$\int_0^T V_\varphi(t) dW_t = V_\varphi(T) W_T - \int_0^T \tilde{V}_\varphi(t) W_t dt$$

puis :

$$\frac{1}{T} \int_0^T [V_\varphi(t) - V_{\varphi^*}(t)] dW_t = \frac{1}{T} [V_\varphi(T) - V_{\varphi^*}(T)] - \frac{1}{T} \int_0^T [\tilde{V}_{\varphi^*}(t)] W_t dt$$

et

$$\begin{aligned} & \sup_{|\varphi - \varphi^*| \leq \eta} \left| \frac{1}{T} \int_0^T [V_\varphi(t) - V_{\varphi^*}(t)] dW_t \right| \\ & \leq \eta \cdot \left\{ \frac{1}{T} U_T \cdot |W_T| + \frac{1}{T} \int_0^T \tilde{U}_t \cdot |W_t| dt \right\}. \end{aligned}$$

Comme :

$$E(U_T \cdot |W_T|) \leq [E(U_T^2)]^{1/2} \cdot [E(W_T^2)]^{1/2} = T^{1/2} [E(U_T^2)]^{1/2}$$

et

$$\begin{aligned} E\left(\int_0^T \tilde{U}_t |W_t| dt\right) & \leq E\left[\left(\int_0^T \tilde{U}_t dt\right) \cdot \left(\sup_{t \leq T} |W_t|\right)\right] \\ & \leq \left(E\left[\left(\int_0^T \tilde{U}_t dt\right)^2\right]\right)^{1/2} \cdot \left(E\left[\sup_{t \leq T} |W_t|^2\right]\right)^{1/2} \\ & \leq C_2^{1/2} T^{1/2} \left[E\left(\int_0^T \tilde{U}_t dt\right)^2\right]^{1/2}, \end{aligned}$$

compte-tenu des hypothèses on a :

$$\sup_{|\varphi - \varphi^*| \leq \eta} \left| \frac{1}{T} \int_0^T [V_\varphi(t) - V_{\varphi^*}(t)] dW_t \right| \leq \eta \cdot M_T$$

où M_T est une variable aléatoire d'espérance bornée indépendamment de T .

Ainsi, traduisant les hypothèses de cette étude particulière, pour les processus $\left\{ \sigma^{-1}(t, \cdot) \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} Q_\varphi(t, \cdot); t \geq 0 \right\}$ nous pouvons décrire des conditions pour lesquelles l'hypothèse (E4) de la proposition I. 2.2.1. est satisfaite.

Dans le chapitre suivant nous rencontrerons une situation où l'hypothèse (E4) sera vérifiée sans que les conditions envisagées ici le soient. De même des hypothèses plus faibles que les hypothèses (K2) et (K3) suffiront à atteindre les conclusions visées par les dites hypothèses.

1.3. Sur la construction des estimateurs.

Contrairement au cas de certains modèles à coefficients dépendant linéairement des paramètres à estimer, nous ne disposons pas ici d'expressions explicites des estimateurs. Toutefois les propositions 1.1.1. et 1.2.2. montrent que l'identification des paramètres d'un modèle, qu'il s'agisse de ceux apparaissant dans le coefficient de diffusion ou de ceux figurant dans le coefficient de dérive, repose sur la résolution d'une équation vectorielle du type :

$$(10) \quad f_{\nu}(\varphi, x) = 0 \quad ; \quad \nu \in \Lambda, \quad \varphi \in \Phi \subset \mathbb{R}^d, \quad x \in C,$$

où pour $\nu \in \Lambda$, f_{ν} est une application de $\Phi \times C$ dans \mathbb{R}^d .

S'il s'agit des paramètres de diffusion on a :

$$\Lambda = \mathbb{N}^* \quad \text{et} \quad f_{\nu}(\varphi, x) = F(\varphi, x) - \gamma^{\nu}(x)$$

où $P_{\varphi^*} - \lim_{\nu \rightarrow +\infty} f_{\nu}(\varphi, \cdot) = 0$

et $f_{\nu}(\cdot, x)$ est supposée différentiable de différentielle $df_{\nu}(\varphi, x) = F^{(1)}(\varphi, x)$ non singulière en φ^* et telle que, pour tout $\nu \in \mathbb{N}^*$,

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} \|df_{\nu}(\varphi, \cdot) - df_{\nu}(\varphi^*, \cdot)\| \leq \varepsilon(\eta)$$

avec $P_{\varphi^*} - \lim_{\eta \rightarrow +\infty} \varepsilon(\eta) = 0$.

S'il s'agit des paramètres de dérive on a :

$$\Lambda = \mathbb{R}_+^* \quad \text{et} \quad f_{\nu}(\varphi, x) = \frac{1}{\nu} L_{\nu}^{(1)}(\varphi, x)$$

où $P_{\varphi^*} - \lim_{\nu \rightarrow +\infty} f_{\nu}(\varphi, \cdot) = 0$

et $f_{\nu}(\cdot, x)$ est supposée différentiable ($df_{\nu}(\varphi, x) = \frac{1}{\nu} L_{\nu}^{(2)}(\varphi, x)$) avec :

$$P_{\varphi^*} - \lim_{\nu \rightarrow +\infty} df_{\nu}(\varphi^*, \cdot) = \Gamma$$

où Γ est une matrice non singulière et :

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} \|df_{\nu}(\varphi, \cdot) - df_{\nu}(\varphi^*, \cdot)\| \leq \varepsilon(\eta, \nu)$$

avec, pour tout $\rho > 0$,

$$\lim_{\eta \rightarrow 0} \limsup_{\nu \rightarrow +\infty} P_{\varphi^*} [\varepsilon(\eta, \nu) \geq \rho] = 0.$$

Le problème se pose donc sensiblement dans les mêmes termes pour les deux cas. Une méthode générale d'approche d'un tel problème est basée sur l'algorithme de NEWTON-RAPHSON (voir par exemple T.L.SAATY [46] ch. 2) :

$$(11) \quad \varphi_{\nu}^{(n+1)}(x) = \varphi_{\nu}^{(n)}(x) - [df_{\nu}(\varphi_{\nu}^{(n)}(x), x)]^{-1} [f_{\nu}(\varphi_{\nu}^{(n)}(x), x)]$$

qui, dans le cas des paramètres de diffusion, s'écrit :

$$(11') \quad \varphi_{\nu}^{(n+1)}(x) = \varphi_{\nu}^{(n)}(x) - [F^{(1)}(\varphi_{\nu}^{(n)}(x), x)]^{-1} [F(\varphi_{\nu}^{(n)}(x), x) - \gamma^{\nu}(x)]$$

et, dans le cas des paramètres de dérive, prend la forme :

$$(11'') \quad \varphi_{\nu}^{(n+1)}(x) = \varphi_{\nu}^{(n)}(x) - [L_{\nu}^{(2)}(\varphi_{\nu}^{(n)}(x), x)]^{-1} [L_{\nu}^{(1)}(\varphi_{\nu}^{(n)}(x), x)] .$$

L'initialisation de l'algorithme repose sur la donnée d'une estimation a priori $\varphi_{\nu}^{(0)}$ du paramètre φ^* ce qui suppose par exemple que l'on connaisse un voisinage de la vraie valeur du paramètre. Nous ne discutons pas ici de conditions autres que celles déjà vérifiées sous lesquelles un tel algorithme fournit une suite $\{\varphi_{\nu}^{(n)}(x) ; n \in \mathbb{N}\}$ convergeant vers une solution $\varphi_{\nu}(x)$ de l'équation (10) qui soit l'unique solution de cette équation dans un voisinage de φ^* . On peut, par exemple, consulter [46], l'analyse du problème dépendant de la forme particulière que peut prendre l'application f_{ν} mise en jeu dans l'équation.

Remarquons d'une part que, dans le cas des paramètres de diffusion, l'observateur dispose a priori du choix de l'application f_{ν} dans la limite des choix possibles de l'application F basés sur la sélection d'un système d'instantants t_1, t_2, \dots, t_k et d'un système de composantes des matrices $\int_0^{t_1} \sigma_{\varphi}(s, x) \cdot \int_0^{t_2} \sigma_{\varphi}(s, x) ds$. D'autre part, dans le cas des paramètres de dérive, l'application f_{ν} est imposée à l'observateur comme étant, au facteur ν^{-1} près, la dérivée de la fonction de vraisemblance. Même dans l'hypothèse où l'observateur disposerait d'une autre méthode lui permettant de construire une famille convergente d'estimateurs $\{\varphi_{\nu}^{(0)} ; \nu > 0\}$, l'approche algorithmique indiquée garde son intérêt au sens suivant : initialisant l'algorithme (11'') par $\varphi_{\nu}^{(0)}(x)$, on peut montrer, comme cela est fait dans [42] ch. I, PHAM DINH TUAN dans le cadre des séries chronolo-

giques, que si $\varphi_{\nu}^{(0)}$ est $\nu^{1/2}$ consistant au sens que $\{\nu^{1/2}(\varphi_{\nu}^{(0)} - \varphi^*) ; \nu > 0\}$ est bornée en probabilité, alors la première itération de l'algorithme fournit un estimateur $\varphi_{\nu}^{(1)} = \varphi_{\nu}^{(0)} - [L_{\nu}^{(2)}(\varphi_{\nu}^{(0)}, \dots)]^{-1} [L_{\nu}^{(1)}(\varphi_{\nu}^{(0)}, \dots)]$ qui est tel que $\{\nu^{1/2}(\varphi_{\nu}^{(1)} - \hat{\varphi}_{\nu}) ; \nu > 0\}$ converge en probabilité vers zéro, où $\hat{\varphi}_{\nu}$ est l'estimateur consistant de maximum de vraisemblance. Ainsi $\varphi_{\nu}^{(1)}$ est un estimateur approchant l'estimateur de maximum de vraisemblance et qui est lui-même asymptotiquement efficace, qualité qu'a priori ne possédait pas $\varphi_{\nu}^{(0)}$.

Dans le cas particulier du modèle que nous envisageons dans le paragraphe suivant K.W. ILIFF donne dans [27] des résultats numériques d'expériences de simulation et sur données réelles qui montrent l'utilité pratique de cette approche.

§ 2. - APPLICATION : ESTIMATION DE PARAMETRES DANS LE MODELE DE FILTRAGE LINEAIRE A COEFFICIENTS CONSTANTS.

Nous considérons un système dynamique stochastique dont le processus d'état (X_t) , régi par une équation différentielle stochastique linéaire à coefficients constants, est observé par une sortie bruitée (Y_t) liée au processus d'état par une équation de même type. Le modèle s'écrit sous la forme :

$$(12) \quad \begin{cases} dX_t = \theta X_t dt + \Lambda dW_t^1 & ; \quad t \geq 0 \\ X_0 = 0 \\ dY_t = F X_t dt + G dW_t^2 & ; \quad t \geq 0 \\ Y_0 = 0 \end{cases}$$

- où * θ, Λ, F, G sont des matrices constantes $n \times n$, $n \times n$, $p \times n$ et $p \times p$ respectivement, partiellement inconnues,
- * et $(W_t^1 ; t \geq 0)$ et $(W_t^2 ; t \geq 0)$ sont des mouvements browniens standards indépendants n -dimensionnel et p -dimensionnel respectivement.

Si φ désigne le vecteur des paramètres inconnus dans les matrices θ, Λ, F et G , le problème est d'estimer la vraie valeur φ^* de φ au vu de l'observation d'une trajectoire de la sortie (Y_t) .

Si l'état (X_t) était aussi directement observable, écrivant le modèle sous la forme :

$$\left\{ \begin{array}{l} d \begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \theta & 0 \\ F & 0 \end{pmatrix} \begin{pmatrix} X_t \\ Y_t \end{pmatrix} dt + \begin{pmatrix} \Lambda & 0 \\ 0 & G \end{pmatrix} d \begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix} ; \quad t \geq 0 \\ \begin{pmatrix} X_0 \\ Y_0 \end{pmatrix} = 0 \end{array} \right.$$

on voit que le problème d'estimation de φ au vu de l'observation de $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ entrerait dans le champ de l'étude des modèles linéaires que nous avons faite dans le chapitre II, § 1.

Ici nous supposons que seule la sortie (Y_t) est observable. Le problème d'estimation dans la version discrétisée du modèle (12) a fait l'objet de nombreuses études qui étendent essentiellement les méthodes classiques d'estimation au vu d'échantillons indépendants au cas d'observations liées (voir par exemple K.J. ASTRÖM et P. EYCKHOFF [5]).

A notre connaissance seuls A.V. BALAKRISHNAN [7] et A. BAGCHI [6] ont abordé le problème d'identification de paramètres d'un tel modèle en temps continu. Dans [7] A.V. BALAKRISHNAN s'est intéressé à ce problème en vue de modéliser le mouvement d'un avion subissant des perturbations aléatoires dues au vent et de contrôler ensuite ce mouvement. Afin d'utiliser la méthode du maximum de vraisemblance il a été amené à faire la restriction essentielle qui consiste à supposer complètement connue la covariance $G \cdot {}^t G$, sans donner de méthode d'estimation préalable des paramètres inconnus y figurant. Dans [6] A. BAGCHI propose une méthode d'estimation simultanée de tous les paramètres basée sur la recherche d'un zéro du gradient d'une modification de la fonction de vraisemblance construite à partir de la donnée d'une estimation a priori de la matrice $G \cdot {}^t G$.

Dans ce travail nous levons la restriction de l'étude de A.V. BALAKRISHNAN en donnant une méthode d'identification de la covariance $G \cdot {}^tG$ sur un intervalle de temps fini sans qu'il soit besoin, comme le suppose A. BAGCHI, de connaître une estimation a priori de la matrice $G \cdot {}^tG$; notre méthode pourrait d'ailleurs permettre de fournir cette estimation a priori. Une fois la matrice $G \cdot {}^tG$ identifiée et supposée connue nous montrons que le problème d'estimation est du type de celui envisagé dans le paragraphe précédent et vérifions que les hypothèses techniques permettant de conclure à l'identifiabilité asymptotique des paramètres sont satisfaites sous réserve de conditions usuelles de stabilité et d'observabilité. Nous retrouvons en particulier les résultats de A.V. BALAKRISHNAN. En outre nous assurons la normalité asymptotique de l'estimateur et donnons une forme commode de la condition d'identifiabilité qui n'est définie dans [7] que de manière implicite : nous montrons que la matrice mise en jeu dans cette condition est solution d'un problème algébrique dont nous explicitons les équations. Les résultats de notre étude ont été annoncés dans [32] A. LE BRETON pour le cas où les processus mis en jeu ainsi que le paramètre à estimer sont scalaires, l'équation d'observation ne contenant plus de paramètres inconnus après l'identification de la variance du bruit d'observation.

2.1. La structure statistique du problème d'estimation.

On suppose que l'on observe sur un intervalle de temps $[0, T]$, $T > 0$, une sortie $y(\cdot) = (y_t; t \in \mathbb{R}_+) \in C_p = C$ liée à l'état non observable $x(\cdot) = (x_t; t \in \mathbb{R}_+) \in C_n$ d'un système dynamique stochastique, où, pour tout entier $k \geq 1$, C_k désigne l'espace des applications continues de \mathbb{R}_+ dans \mathbb{R}^k , et n et p sont deux entiers tels que $1 \leq p \leq n$.

Φ étant un sous-ensemble de \mathbb{R}^d , $d \geq 1$, θ, Λ et F désignent des applications définies sur Φ (dit ensemble des valeurs du paramètre) à valeurs respectivement dans $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$, $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^n)$ et $\mathcal{L}(\mathbb{R}^n; \mathbb{R}^p)$. On pourra supposer dans la suite que $\varphi \in \Phi$ représente une valeur possible d'une partie fixe des coefficients des matrices θ, Λ et F , l'autre partie étant donnée; $\theta(\varphi)$, $\Lambda(\varphi)$ et $F(\varphi)$ sont alors les valeurs

prises par ces matrices pour la valeur φ de ce système de coefficients. En particulier θ, Λ et F seront supposées dérivables à dérivées bornées sur tout borné. On suppose qu'il existe un élément φ^* , inconnu dans Φ , dit "vraie valeur du paramètre φ " et une matrice inconnue G^* non singulière de $\mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$ tels que pour $\varphi = \varphi^*$ et $G = G^*$, $(x(\cdot), y(\cdot))$ soit une trajectoire d'un processus $(X_t, Y_t; t \geq 0)$ solution du système d'équations différentielles stochastiques :

$$(13) \quad \begin{cases} dX_t = \theta(\varphi)X_t dt + \Lambda(\varphi)dW_t^1 & ; \quad t \geq 0, X_0 = 0 \\ dY_t = F(\varphi)X_t dt + GdW_t^2 & ; \quad t \geq 0, Y_0 = 0 \end{cases}$$

où $(W_t^1; t \geq 0)$ et $(W_t^2; t \geq 0)$ sont des processus de mouvements browniens indépendants dans \mathbb{R}^n et \mathbb{R}^p respectivement.

On entend ici par solution du système (13) la donnée :

- d'un espace probabilisé (Ω, \mathcal{A}, P) muni d'une famille croissante $\mathcal{A}(\cdot) = (\mathcal{A}_t; t \geq 0)$ de sous-tribus de \mathcal{A} ,
- de deux mouvements browniens standards $(W_t^1; t \geq 0)$ et $(W_t^2; t \geq 0)$ dans \mathbb{R}^n et \mathbb{R}^p respectivement, relativement à $(\Omega, \mathcal{A}(\cdot), P)$, qui soient P -indépendants,
- de deux fonctions aléatoires $X : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^n$ et $Y : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}^p$ continues et adaptées à $\mathcal{A}(\cdot)$ telles que :

$$\begin{cases} X_t = \int_0^t \theta(\varphi)X_s ds + \Lambda(\varphi)W_t^1 & ; \quad t \geq 0 \\ Y_t = \int_0^t F(\varphi)X_s ds + GW_t^2 & ; \quad t \geq 0 . \end{cases}$$

Pour φ fixé dans Φ et G donnée dans $\mathcal{L}(\mathbb{R}^p; \mathbb{R}^p)$, on est assuré (théorème I. 1.3.1) de l'existence et de l'unicité en loi de la solution de (13). Cette loi, qui ne dépend d'ailleurs que de $(\theta(\varphi), \Lambda(\varphi), \Lambda(\varphi), F(\varphi), G, {}^tG)$, peut être définie comme étant la mesure induite sur C_{n+p} par la solution particulière suivante de (13) :

$$\begin{aligned} \Omega &= C_{n+p} = C_n \times C_p & ; & \quad \mathcal{A} = S^{n+p} = S^n \otimes S^p & ; & \quad \mathcal{A}_t = S_t^{n+p} = S_t^n \otimes S_t^p, \quad t \geq 0 \\ P &= \mu_*^{n+p} = \mu_*^n \otimes \mu_*^p \end{aligned}$$

où, pour tout entier $k \geq 1$, S_t^k , Π_t^k , μ_*^k et $\mu_*^{k,T}$ désignent les analogues, relativement à C_k , de $S_t^p = S_t$, $\Pi_t^p = \Pi_t$, $\mu_*^p = \mu_*$ et $\mu_*^{p,T} = \mu_*^T$ relativement à $C_p = C$,

$$W_t^1 = \tilde{\Pi}_t^n = \Pi_t^n \circ \zeta_1 \quad \text{et} \quad W_t^2 = \tilde{\Pi}_t^p = \Pi_t^p \circ \zeta_2, \quad t \geq 0,$$

où ζ_1 et ζ_2 désignent les projections de $C_n \times C_p$ sur C_n et C_p respectivement, $(X_t, Y_t; t \geq 0)$ est la solution de Itô (unique à l'indistinguabilité près) du système (13) relativement à la base de processus $(\Omega, \mathcal{A}, \mathcal{A}(\cdot), \mu_*^{n+p})$.

On est alors amené naturellement à poser le problème d'estimation du paramètre $(\varphi, G, {}^tG)$, au vu de l'observation de la sortie $y(\cdot)$ sur l'intervalle de temps $[0, T]$, dans la structure statistique :

$$(C, S_T, \{\mu_{\varphi, G}^{p,T}; \varphi \in \Phi, G \in \mathcal{L}^*(\mathbb{R}^p; \mathbb{R}^p)\})$$

où, pour tout $\varphi \in \Phi$ et $G \in \mathcal{L}^*(\mathbb{R}^p; \mathbb{R}^p)$, $\mu_{\varphi, G}^{p,T}$ (resp. $\mu_\varphi^{n,T}$) désigne la seconde (resp. première) marginale de la mesure $\mu_{\varphi, G}^T$ restriction à S_T^{n+p} de la mesure $\mu_{\varphi, G}$ induite de μ_*^{n+p} par le processus $(X_t^{\varphi, G}, Y_t^{\varphi, G}; t \geq 0)$ solution de (13) lorsque les paramètres sont φ et G .

On étudie maintenant la famille des mesures

$\{\mu_{\varphi, G}^{p,T}; \varphi \in \Phi, G \in \mathcal{L}^*(\mathbb{R}^p; \mathbb{R}^p)\}$ et on montre en particulier que, pour tout $\varphi \in \Phi$ et $G \in \mathcal{L}^*(\mathbb{R}^p; \mathbb{R}^p)$, la mesure $\mu_{\varphi, G}^{p,T}$ est équivalente à la mesure $\mu_G^{p,T}$ où si μ_G^p désigne la mesure de Wiener de covariance $G \cdot {}^tG$ sur C_p , $\mu_G^{p,T} = \mu_G^p / S_T$.

2.1.1. LEMME : "Les mesures $\mu_{\varphi, G}^T$ et $\mu_\varphi^{n,T} \otimes \mu_G^{p,T}$ sur S_T^{n+p} sont équivalentes, la dérivée de Radon-Nikodym étant donnée par :

$$(14) \quad \frac{d\mu_{\varphi, G}^T}{d(\mu_\varphi^{n,T} \otimes \mu_G^{p,T})} = \exp\left\{ \int_0^T \langle F(\varphi) \tilde{\Pi}_t^n, (G \cdot {}^tG)^{-1} d\tilde{\Pi}_t^p \rangle - \frac{1}{2} \int_0^T \|G^{-1} F(\varphi) \tilde{\Pi}_t^n\|^2 dt \right\}$$

où l'intégrale stochastique est définie relativement à la mesure $\mu_\varphi^{n,T} \otimes \mu_G^{p,T}$.

Démonstration : Considérons le processus $(Z_t = \begin{pmatrix} U_t \\ V_t \end{pmatrix} ; t \geq 0)$ défini relativement à $(C^{n+p}, S^{n+p}, (S_t^{n+p})_{t \geq 0}, \mu_*^{n+p})$ par :

$$dZ_t = \begin{pmatrix} \theta(\varphi) & 0 \\ 0 & 0 \end{pmatrix} Z_t dt + \begin{pmatrix} \Lambda(\varphi) & 0 \\ 0 & G \end{pmatrix} d \begin{pmatrix} \tilde{\Pi}_t^n \\ \tilde{\Pi}_t^p \end{pmatrix} ; \quad t \geq 0, \quad Z_0 = 0,$$

et la mesure $\tilde{\mu}^T$ sur S_T^{n+p} , absolument continue par rapport à $\mu_*^{n+p, T}$, avec pour dérivée de Radon-Nikodym :

$$\begin{aligned} \frac{d\tilde{\mu}^T}{d\mu_*^{n+p, T}} &= \exp\left\{ \int_0^T \left\langle \begin{pmatrix} 0 & 0 \\ 0 & G^{-1} \end{pmatrix} \begin{pmatrix} \theta(\varphi) & 0 \\ F(\varphi) & 0 \end{pmatrix} Z_t, d \begin{pmatrix} \tilde{\Pi}_t^n \\ \tilde{\Pi}_t^p \end{pmatrix} \right\rangle - \frac{1}{2} \int_0^T \left\| \begin{pmatrix} 0 & 0 \\ 0 & G^{-1} \end{pmatrix} \begin{pmatrix} \theta(\varphi) & 0 \\ F(\varphi) & 0 \end{pmatrix} Z_t \right\|^2 dt \right\} \\ &= \exp\left\{ \int_0^T \langle G^{-1} F(\varphi) U_t, d\tilde{\Pi}_t^p \rangle - \frac{1}{2} \int_0^T \|G^{-1} F(\varphi) U_t\|^2 dt \right\}. \end{aligned}$$

Les théorèmes I. 1.3.3 et I. 1.2.1 assurent que $\tilde{\mu}^T$ est une probabilité et que de plus le processus $\{\check{\Pi}_t ; t \in [0, T]\}$ défini par :

$$\check{\Pi}_t = \begin{pmatrix} \tilde{\Pi}_t^n \\ \tilde{\Pi}_t^p \end{pmatrix} - \int_0^t \begin{pmatrix} 0 & 0 \\ 0 & G^{-1} \end{pmatrix} \begin{pmatrix} \theta(\varphi) & 0 \\ F(\varphi) & 0 \end{pmatrix} Z_s ds = \begin{pmatrix} \tilde{\Pi}_t^n \\ \tilde{\Pi}_t^p - \int_0^t G^{-1} F(\varphi) U_s ds \end{pmatrix}$$

est un processus de Wiener sur la base de processus $(C^{n+p}, (S_t^{n+p})_{0 \leq t \leq T}, \tilde{\mu}^T)$ et :

$$Z_t = \int_0^t \begin{pmatrix} \theta(\varphi) & 0 \\ F(\varphi) & 0 \end{pmatrix} Z_s ds + \begin{pmatrix} \Lambda(\varphi) & 0 \\ 0 & G \end{pmatrix} d\check{\Pi}_t ; \quad t \in [0, T].$$

D'après le théorème I. 1.2.3 la loi de $(Z_t ; t \in [0, T] ; \tilde{\mu}^T)$ est équivalente à celle de $(Z_t ; t \in [0, T] ; \mu_*^{n+p, T})$ la dérivée de Radon-Nikodym s'écrivant :

$$\exp\left\{ \int_0^T \langle G^{-1} F(\varphi) \tilde{\Pi}_t^n, G^{-1} d\tilde{\Pi}_t^p \rangle - \frac{1}{2} \int_0^T \|G^{-1} F(\varphi) \tilde{\Pi}_t^n\|^2 dt \right\}$$

où l'intégrale stochastique est définie relativement à cette dernière loi qui n'est autre que $\mu_\varphi^{n, T} \otimes \mu_G^{p, T}$.

Le résultat du lemme est donc acquis compte-tenu de l'unicité en loi de la solution des équations :

$$Z_t = \int_0^t \begin{pmatrix} \theta(\varphi) & 0 \\ F(\varphi) & 0 \end{pmatrix} Z_s ds + \begin{pmatrix} \Lambda(\varphi) & 0 \\ 0 & G \end{pmatrix} \text{ (mouvement brownien) } ; \quad t \geq 0$$

et
$$Z_t = \int_0^t \begin{pmatrix} \theta(\varphi) & 0 \\ 0 & 0 \end{pmatrix} Z_s ds + \begin{pmatrix} \Lambda(\varphi) & 0 \\ 0 & G \end{pmatrix} \text{ (mouvement brownien) ; } t \geq 0 \text{ . } \blacksquare$$

On est alors en mesure de démontrer la propriété annoncée de la famille de mesures $\mu_{\varphi, G}^{p, T}$; $\varphi \in \Phi$, $\mathfrak{L}^*(\mathbb{R}^p; \mathbb{R}^p)$ } décrite dans l'énoncé suivant qui, sous une forme adaptée à notre problème, fait la synthèse d'un résultat contenu dans [17] T.E. DUNCAN et des résultats classiques de la théorie du filtrage linéaire (voir par exemple R.S. BUCY et P.D. JOSEPH [14]).

2.1.2. PROPOSITION : "Pour tout $\varphi \in \Phi$ et tout $G \in \mathfrak{L}^*(\mathbb{R}^p; \mathbb{R}^p)$, la mesure $\mu_{\varphi, G}^{p, T}$ est équivalente à la mesure $\mu_G^{p, T}$ avec pour dérivée de Radon-Nikodym :

$$(15) \quad \frac{d\mu_{\varphi, G}^{p, T}}{d\mu_G^{p, T}} = \exp\left\{ \int_0^T \langle F(\varphi) \alpha_{\varphi}^G(t, \cdot), (G \cdot {}^t G)^{-1} d\Pi_t^p \rangle - \frac{1}{2} \int_0^T \|G^{-1} F(\varphi) \alpha_{\varphi}^G(t, \cdot)\|^2 dt \right\}$$

où l'intégrale stochastique est définie relativement à la base de processus $(C_p, (S_t^p)_{t \geq 0}, \mu_G^p)$, la fonction aléatoire $(\alpha_{\varphi}^G(t, \cdot) ; t \geq 0)$ étant définie sur C_p par :

$$(16) \quad \alpha_{\varphi}^G(t, y) = Q_{\varphi}^G(t) \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} y_t - R_{\varphi}^G(t) \int_0^t [R_{\varphi}^G(s)]^{-1} [Q_{\varphi}^G(s) \cdot {}^t \theta(\varphi) + \Lambda(\varphi) \cdot {}^t \Lambda(\varphi)] \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} y_s ds$$

pour $y = (y_t ; t \geq 0) \in C_p$, $t \geq 0$, avec :

$$(17) \quad \begin{cases} \frac{dQ_{\varphi}^G}{dt}(t) = \theta(\varphi) Q_{\varphi}^G(t) + Q_{\varphi}^G(t) \cdot {}^t \theta(\varphi) + \Lambda(\varphi) \cdot {}^t \Lambda(\varphi) - Q_{\varphi}^G(t) \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi) Q_{\varphi}^G(t) \\ Q_{\varphi}^G(0) = 0 \end{cases} \quad t \geq 0$$

et :

$$(18) \quad \begin{cases} \frac{dR_{\varphi}^G}{dt}(t) = [\theta(\varphi) - Q_{\varphi}^G(t) \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi)] R_{\varphi}^G(t) ; \quad t \geq 0 \\ R_{\varphi}^G(0) = \mathbb{1}_n \text{ (matrice identité de } \mathfrak{L}(\mathbb{R}^n; \mathbb{R}^n)) \text{ .} \end{cases}$$

Démonstration : La mesure $\mu_{\varphi, G}^T$ étant d'après le lemme 2.1.1. équivalente à la mesure produit $\mu_{\varphi}^{n, T} \otimes \mu_G^{p, T}$, on sait que la mesure $\mu_{\varphi, G}^{p, T}$ est équivalente à la mesure $\mu_G^{p, T}$, la dérivée de Radon-Nikodym étant donnée par :

$$\frac{d\mu_{\varphi, G}^{p, T}}{d\mu_G^{p, T}} = E_{\mu_{\varphi}^{n, T}} \left[\frac{d\mu_{\varphi, G}^T}{d(\mu_{\varphi}^{n, T} \otimes \mu_G^{p, T})} \right], \quad \mu_G^{p, T} \text{ presque sûrement.}$$

Dans [17], T.E. DUNCAN montre (théorème 2) que, dans le problème de filtrage de $\tilde{\pi}_t^n$ au vu de $\tilde{\pi}_s^p$, $0 \leq s \leq t$, relativement à la base $(C_{n+p}, (S_t^{n+p})_{t \geq 0}, \mu_{\varphi, G})$, on a :

$$E_{\mu_{\varphi, G}} [\tilde{\pi}_t^n / \tilde{\pi}_s^p ; 0 \leq s \leq t] = \hat{\pi}_t^n = \frac{E_{\mu_{\varphi}^{n, T}} [\tilde{\Phi}_t^{\varphi} \tilde{\pi}_t^n]}{E_{\mu_{\varphi}^{n, T}} [\tilde{\Phi}_t^{\varphi}]}$$

où $\tilde{\Phi}_t^{\varphi} = \exp \left[\int_0^t \langle F(\varphi) \tilde{\pi}_s^n, (G \cdot {}^t G)^{-1} d\tilde{\pi}_s^p \rangle - \frac{1}{2} \int_0^t \|G^{-1} F(\varphi) \tilde{\pi}_s^n\|^2 ds \right]$

soit, d'après le lemme 2.1.1.,

$$\tilde{\Phi}_t^{\varphi} = \frac{d\mu_{\varphi, G}^t}{d(\mu_{\varphi}^{n, t} \otimes \mu_G^{p, t})}$$

T.E. DUNCAN montre aussi (théorème 3) que :

$$d \text{Log } E_{\mu_{\varphi}^{n, T}} (\tilde{\Phi}_t^{\varphi}) = \langle F(\varphi) \hat{\pi}_t^n, (G \cdot {}^t G)^{-1} d\tilde{\pi}_t^p \rangle - \frac{1}{2} \|G^{-1} F(\varphi) \hat{\pi}_t^n\|^2 dt$$

Alors, on a :

$$(19) \quad \frac{d\mu_{\varphi, G}^{p, T}}{d\mu_G^{p, T}} = \exp \left\{ \int_0^T \langle F(\varphi) \hat{\pi}_t^n, (G \cdot {}^t G)^{-1} d\tilde{\pi}_t^p \rangle - \frac{1}{2} \int_0^T \|G^{-1} F(\varphi) \hat{\pi}_t^n\|^2 dt \right\}$$

Or, relativement à la base de processus $(C_{n+p}, (S_t^{n+p})_{t \geq 0}, \mu_{\varphi, G})$ on a :

$$\begin{cases} d\tilde{\pi}_t^n = \theta(\varphi) \tilde{\pi}_t^n dt + \Lambda(\varphi) dB_t^1 & ; \quad t \geq 0 ; \tilde{\pi}_0^n = 0 \\ d\tilde{\pi}_t^p = F(\varphi) \tilde{\pi}_t^n dt + G dB_t^2 & ; \quad t \geq 0 ; \tilde{\pi}_0^p = 0 \end{cases}$$

où $(B_t^1 ; t \geq 0)$ et $(B_t^2 ; t \geq 0)$ sont deux mouvements browniens indépendants.

Les résultats classiques de la théorie du filtrage linéaire [14] permettent alors d'assurer que le processus $(\hat{\Pi}_t^n ; t \geq 0)$ admet la représentation :

$$\begin{cases} d\hat{\Pi}_t^n = \theta(\varphi)\hat{\Pi}_t^n dt + Q_\varphi^G(t) \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1} (d\tilde{\Pi}_t^p - F(\varphi)\hat{\Pi}_t^n dt) \\ \hat{\Pi}_0^n = 0 \end{cases}$$

où $(Q_\varphi^G(t) ; t \geq 0)$ est donné par (17) et, relativement à $\mu_{\varphi, G}$, le processus :

$$(G^{-1}(\tilde{\Pi}_t^p - \int_0^t F(\varphi)\hat{\Pi}_s^n ds) ; t \geq 0)$$

est un mouvement brownien standard appelé "processus d'innovation" par T. KAILATH [29]. On a donc, $(R_\varphi^G(t) ; t \geq 0)$ étant donné par (18),

$$\hat{\Pi}_t^n = \int_0^t R_\varphi^G(t) [R_\varphi^G(s)]^{-1} Q_\varphi^G(s) \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1} d\tilde{\Pi}_s^p$$

et, intégrant par parties, il vient :

$$\hat{\Pi}_t^n = Q_\varphi^G(t) \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1} \tilde{\Pi}_t^p - R_\varphi^G(t) \int_0^t \frac{d}{ds} ([R_\varphi^G(s)]^{-1} Q_\varphi^G(s)) \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1} \tilde{\Pi}_s^p ds$$

soit, d'après (17) et (18),

$$\begin{aligned} \hat{\Pi}_t^n &= Q_\varphi^G(t) \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1} \tilde{\Pi}_t^p \\ &\quad - R_\varphi^G(t) \int_0^t [R_\varphi^G(s)]^{-1} \cdot [Q_\varphi^G(s) \cdot {}^t\theta(\varphi) + \Lambda(\varphi) \cdot {}^t\Lambda(\varphi)] \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1} \tilde{\Pi}_s^p ds . \end{aligned}$$

On en déduit immédiatement, compte-tenu de (19), qu'une version de

$$\frac{d\mu_{\varphi, G}^{p, T}}{d\mu_G^{p, T}} \text{ est donnée par (15). } \blacksquare$$

2.1.3. Remarques :

a) Relativement à la base de processus $(C_p, (S_t^p)_{t \geq 0}, \mu_{\varphi, G}^p)$ le processus :

$$(G^{-1}(\tilde{\Pi}_t^p - \int_0^t F(\varphi)\alpha_\varphi^G(s, \cdot) ds) ; t \geq 0)$$

est un mouvement brownien standard.

b) La structure statistique du problème est la même que celle qui concerne l'estimation des paramètres φ et $G \cdot {}^tG$ dans l'équation

différentielle stochastique p-dimensionnelle :

$$\begin{cases} dY_t = F(\varphi) \alpha_{\varphi}^G(t, Y) dt + G dW_t & ; \quad t \geq 0 \\ Y_0 = 0 \end{cases}$$

où pour tout (φ, G) , $\alpha_{\varphi}^G(t, \cdot)$ est définie par (16), (17) et (18), au vu de l'observation d'une trajectoire d'un processus solution. Il est clair que bien que le paramètre φ apparaisse linéairement dans $\theta(\varphi)$, $\Lambda(\varphi)$ et $F(\varphi)$, le coefficient de dérive de cette équation dépend non linéairement de ce paramètre. Ce modèle est donc du domaine de l'étude du premier paragraphe de ce chapitre. Dans la suite nous supprimons les indices p et reprenons les notations que nous avons utilisées précédemment.

2.2. Estimation de la covariance du bruit d'observation.

Nous nous proposons de montrer que notre approche du problème d'estimation des paramètres de diffusion d'un modèle différentiel stochastique nous permet ici d'estimer la matrice de covariance du bruit d'observation dans le modèle (12) sur tout intervalle de temps fini.

Soit, pour tout $N \geq 1$, $0 = t_0^N < t_1^N < \dots < t_{K_N}^N = t$ une subdivision de l'intervalle $[0, t]$ telle que, si $\delta_N = \text{Max}_{k=1, \dots, K_N} |t_k^N - t_{k-1}^N|$, $\lim_{N \rightarrow +\infty} \delta_N = 0$.

On considère la suite de statistiques $\{\hat{\sigma}_t^N ; N \geq 1\}$ définies sur la structure statistique :

$$(C, S_T, \{\mu_{\varphi, G}^T ; \varphi \in \Phi, G \in \mathcal{L}^*(\mathbb{R}^p; \mathbb{R}^p)\})$$

pour $T \geq t > 0$ par :

$$(20) \quad \hat{\sigma}_t^N = \frac{1}{t} \sum_{i=1}^{K_N} (\Pi_{t_k^N} - \Pi_{t_{k-1}^N})^{\otimes 2} ; \quad N \geq 1 .$$

On déduit immédiatement du lemme II. 1.2.2. et de la proposition 2.1.2. l'énoncé suivant :

2.2.1. PROPOSITION : "Dans le modèle (12), la covariance du bruit d'observation est fortement identifiable en probabilité sur tout intervalle de temps fini $[0, t]$ par la suite de statistiques $\{\hat{\sigma}_t^N ; N \geq 1\}$ définie par (20). Si en outre, la série de terme général δ_N est convergente, alors la covariance du bruit d'observation est fortement identifiable presque sûrement par cette suite de statistiques."

Par une démonstration analogue à celle du lemme II. 2.2.3. nous pouvons démontrer le résultat complémentaire suivant qui précise la qualité de l'estimateur $\hat{\sigma}_t^N$:

2.2.2. PROPOSITION : "La suite de variables aléatoires $\{\delta_N^{-1/2}(\hat{\sigma}_t^N - G \cdot {}^tG) ; N \geq 1\}$ est bornée en probabilité pour toute loi $\mu_{\varphi, G}^T$, $\varphi \in \Phi$."

2.2.3. Remarques :

a) Ainsi la matrice de covariance du bruit d'observation peut être identifiée (éventuellement calculée avec probabilité 1) sur tout intervalle de temps fini et il est légitime de la supposer connue pour étudier le problème d'estimation du paramètre φ .

b) Nous sommes ici dans les conditions de la remarque II. 3.1.3.c) où le paramètre $G \cdot {}^tG$ apparaît simultanément dans le coefficient de diffusion et dans le coefficient de dérive. Le problème initial d'estimation des paramètres de dérive est relatif à un paramètre de dimension $d + p^2$. Après estimation de la covariance du bruit d'observation il est relatif au seul paramètre φ de dimension d . Ainsi, en plus du fait qu'elle permet d'appliquer la méthode du maximum de vraisemblance pour l'estimation de φ , notre démarche présente l'avantage sur celle de A. BAGCHI de diminuer la dimension du paramètre à estimer.

2.3. Estimation des paramètres ne figurant pas dans la covariance du bruit d'observation.

Nous supposons désormais que la matrice $G \cdot {}^tG$ est complète-

ment connue, le seul paramètre à estimer étant φ . La structure statistique du problème :

$$(C, S_T, \{\mu_\varphi^T ; \varphi \in \Phi\})$$

est alors dominée par la mesure μ_G^T , avec pour fonction de Log-vraisemblance :

$$(21) \quad L_T(\varphi, \cdot) = \int_0^T \langle F(\varphi) \alpha_\varphi(t, \cdot), (G \cdot {}^t G)^{-1} d\Pi_t \rangle - \frac{1}{2} \int_0^T \|G^{-1} F(\varphi) \alpha_\varphi(t, \cdot)\|^2 dt$$

où la fonction aléatoire $(\alpha_\varphi(t, \cdot) ; t \geq 0)$ est définie sur C par les équations (16), (17) et (18) où nous n'indiquons plus la dépendance de α_φ , Q_φ et R_φ par rapport à G .

Nous assurons d'abord l'existence d'une version régulière de la fonction de Log-vraisemblance (21) :

2.3.1. LEMME : "Pour tout $(t, x) \in \mathbb{R}_+ \times C$, l'application $\varphi \rightsquigarrow b_\varphi(t, x) = F(\varphi) \alpha_\varphi(t, x)$ est trois fois dérivable de dérivées successives $b_\varphi^{(k)}(t, x)$, $k = \overline{1, 3}$. De plus, pour tout $K > 0$, il existe un processus $(U_K(t) ; t \geq 0)$ basé sur C tel que, r désignant un réel, $r \geq 2$, $r > d$,

$$\int_0^T E_G(|U_K(t)|^r) dt < +\infty, \quad T > 0$$

et $\sup_{\|\varphi\| \leq K} \|b_\varphi^{(k)}(t, x)\|^2 \leq U_K(t) ; (t, x) \in \mathbb{R}_+ \times C, k = \overline{1, 3} ."$

La démonstration, très technique, sera omise ici. Nous en donnons les grandes lignes : les propriétés classiques de dérivabilité (par rapport au paramètre) de la solution d'une équation différentielle ordinaire dont les coefficients dépendent d'un paramètre (voir par exemple [52] II. 4. D, J. WARGA) permettent d'assurer la dérivabilité des applications $\varphi \rightsquigarrow (Q_\varphi(t) ; t \in [0, T])$ et $\varphi \rightsquigarrow (R_\varphi(t) ; t \in [0, T])$, les dérivées vérifiant des équations différentielles obtenues par dérivation formelle de (17) et (18). Alors les propriétés de dérivabilité (par rapport au paramètre) de la solution d'une équation intégrale à coefficients dépendant d'un paramètre ([52] II.5.D)

permettent de vérifier que l'application $\varphi \rightsquigarrow (\mathbf{a}_\varphi(t, \mathbf{x}) ; t \in [0, T])$ est dérivable, les dérivées vérifiant des équations obtenues par dérivation formelle de (16). Par suite l'application $\varphi \rightsquigarrow (\mathbf{b}_\varphi(t, \mathbf{x}) ; t \in [0, T])$ est elle aussi dérivable.

L'existence d'un processus $(U_K(t) ; t \geq 0)$ vérifiant les conditions du lemme est une conséquence facile de ce que les dérivées successives de θ, Λ et F sont bornées sur tout borné et du fait que, relativement à $\mu_G, (\Pi_t ; t \geq 0)$ est un processus gaussien.

Alors, d'après le lemme 1.2.1., il existe des fonctions aléatoires $\varphi \rightsquigarrow \hat{L}_T(\varphi, \cdot), \varphi \rightsquigarrow \hat{L}_T^{(1)}(\varphi, \cdot) = (\hat{L}_{T,i}^{(1)}(\varphi, \cdot) ; i = \overline{1, d})$ et $\varphi \rightsquigarrow \hat{L}_T^{(2)}(\varphi, \cdot) = (\hat{L}_{T,i,j}^{(2)}(\varphi, \cdot) ; i, j = \overline{1, d})$ à trajectoires continues telles que l'on ait :

(i) pour tout $\varphi \in \Phi, \mu_G$ -presque sûrement :

$$\hat{L}_T(\varphi, \cdot) = L_T(\varphi, \cdot) \quad \text{et} \quad \hat{L}_T^{(k)}(\varphi, \cdot) = L_T^{(k)}(\varphi, \cdot), \quad k = \overline{1, 2}$$

(ii) pour $i, j = \overline{1, d}$ et $\mathbf{x} \in C$:

$$\hat{L}_{T,i}^{(1)}(\varphi, \mathbf{x}) = \frac{\partial}{\partial \varphi_i} \hat{L}_T(\varphi, \mathbf{x}) \quad \text{et} \quad \hat{L}_{T,i,j}^{(2)}(\varphi, \mathbf{x}) = \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} \hat{L}_T(\varphi, \mathbf{x}),$$

où :

$$(22) \quad L_{T,i}^{(1)}(\varphi, \cdot) = \int_0^T \left\langle \frac{\partial}{\partial \varphi_i} (F(\varphi) \mathbf{a}_\varphi(t)), (G \cdot {}^t G)^{-1} (d\Pi_t - F(\varphi) \mathbf{a}_\varphi(t) dt) \right\rangle ; \quad i = \overline{1, d}$$

$$(23) \quad L_{T,i}^{(2)}(\varphi, \cdot) = \int_0^T \left\langle \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} (F(\varphi) \mathbf{a}_\varphi(t)), (G \cdot {}^t G)^{-1} (d\Pi_t - F(\varphi) \mathbf{a}_\varphi(t) dt) \right\rangle \\ - \int_0^T \left\langle \frac{\partial}{\partial \varphi_i} (F(\varphi) \mathbf{a}_\varphi(t)), (G \cdot {}^t G)^{-1} \frac{\partial}{\partial \varphi_j} (F(\varphi) \mathbf{a}_\varphi(t)) \right\rangle dt ; \quad i, j = \overline{1, d}.$$

Nous faisons désormais les hypothèses suivantes, $\mathcal{V}(\varphi^*)$ désignant un voisinage de φ^* :

(L1) Pour tout $\varphi \in \mathcal{V}(\varphi^*)$, la matrice $\theta(\varphi)$ est stable i.e. a des valeurs propres dont les parties réelles sont toutes strictement négatives.

(L2) Pour tout $\varphi \in \mathcal{V}(\varphi^*)$, le couple $(F(\varphi), \theta(\varphi))$ est complètement observable i.e. la matrice :

$$\int_0^{+\infty} e^{s^t \theta(\varphi)} \cdot {}^t F(\varphi) \cdot (G \cdot {}^t G)^{-1} F(\varphi) e^{s \theta(\varphi)} ds \text{ est non singulière.}$$

On sait (voir par exemple [14] ch. V, R.S. BUCY et P.D. JOSEPH) que lorsque (L1) et (L2) sont vérifiées, alors, si $(Q_\varphi(t); t \geq 0)$ est donnée par (17), on a :

$$\lim_{t \rightarrow +\infty} Q_\varphi(t) = Q_\varphi^\infty$$

où Q_φ^∞ est l'unique solution dans $\mathcal{L}^*(\mathbb{R}^n; \mathbb{R}^n)$ de l'équation :

$$(24) \quad \theta(\varphi) \cdot Q + Q \cdot {}^t \theta(\varphi) + \Lambda(\varphi) \cdot {}^t \Lambda(\varphi) - Q \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi) Q = 0$$

et que de plus la matrice $[\theta(\varphi) - Q_\varphi^\infty \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi)]$ est stable. On énonce alors un lemme précisant l'approximation de $Q_\varphi(t)$ par Q_φ^∞ :

2.3.2. LEMME : "Pour tout $\varphi \in \mathcal{V}(\varphi^*)$ on a, pour tout $t > 0$:

$$\|Q_\varphi(t) - Q_\varphi^\infty\| \leq K_\varphi e^{-\alpha(\varphi)t}$$

$$\left\| \frac{\partial}{\partial \varphi_i} [Q_\varphi(t) - Q_\varphi^\infty] \right\| \leq K_\varphi e^{-\alpha(\varphi)t}, \quad i = \overline{1, d},$$

$$\left\| \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} [Q_\varphi(t) - Q_\varphi^\infty] \right\| \leq K_\varphi e^{-\alpha(\varphi)t}, \quad i, j = \overline{1, d},$$

où K_φ et $\alpha(\varphi)$ sont des constantes strictement positives indépendantes de t ."

Démonstration : Remarquons que la matrice Q_φ^∞ peut être définie comme étant la solution de l'équation différentielle (17) pour la condition initiale Q_φ^∞ . On écrit alors, posant $Q_\varphi^\infty - Q_\varphi(t) = \Delta_\varphi(t)$:

$$\begin{cases} \Delta_\varphi(0) = Q_\varphi^\infty \\ \frac{d}{dt} \Delta_\varphi(t) = [\theta(\varphi) - Q_\varphi(t) \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi)] \Delta_\varphi(t) + \Delta_\varphi(t) \cdot {}^t [\theta(\varphi) - Q_\varphi^\infty \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi)] \end{cases}$$

Donc si $R_\varphi(t)$ (resp ${}^t R_\varphi^\infty(t)$) est la résolvante de $(\theta(\varphi) - Q_\varphi(t) \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi))$

(resp^t $(\theta(\varphi) - Q_\varphi^\infty \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1}F(\varphi))$, on a :

$$\Delta_\varphi(t) = R_\varphi(t)Q_\varphi^\infty \cdot {}^tR_\varphi^\infty(t) .$$

Alors, compte-tenu des hypothèses (L1) et (L2) et du théorème 5.4. de [14], il vient :

$$\|\Delta_\varphi(t)\| \leq C\|Q_\varphi^\infty\|e^{-\alpha(\varphi)t}$$

où C et $\alpha(\varphi)$ sont des constantes strictement positives indépendantes de t . L'assertion du lemme est donc démontrée pour la différence $(Q_\varphi(t) - Q_\varphi^\infty)$.

Ecrivant les dérivées partielles $\frac{\partial}{\partial\varphi_i}(Q_\varphi(t) - Q_\varphi^\infty)$ comme solution d'une équation différentielle faisant apparaître $\Delta_\varphi(t)$, en utilisant la majoration précédente il est facile d'obtenir une majoration analogue de $\|\frac{\partial}{\partial\varphi_i}(Q_\varphi(t) - Q_\varphi^\infty)\|$.

De même les dérivées partielles secondes $\frac{\partial^2}{\partial\varphi_i\partial\varphi_j}(Q_\varphi(t) - Q_\varphi^\infty)$ sont solutions d'une équation différentielle contenant $\Delta_\varphi(t)$ et $\frac{\partial}{\partial\varphi_j}(Q_\varphi(t) - Q_\varphi^\infty)$, et utilisant les majorations déjà obtenues de ces expressions on peut montrer une majoration de même nature de $\|\frac{\partial^2}{\partial\varphi_i\partial\varphi_j}(Q_\varphi(t) - Q_\varphi^\infty)\|$. ■

Remarquons que, relativement à la mesure μ_{φ^*} , les processus $(a_{\varphi^*}(t, \cdot); t \geq 0)$, $(\frac{\partial}{\partial\varphi_i} a_{\varphi^*}(t, \cdot); t \geq 0)$, $i = \overline{1, d}$ et $(\frac{\partial^2}{\partial\varphi_i\partial\varphi_j} a_{\varphi^*}(t, \cdot); t \geq 0)$ peuvent être définis respectivement par :

$$(24) \quad a_{\varphi^*}(t, \cdot) = \int_0^t \theta(\varphi^*) a_{\varphi^*}(s, \cdot) ds + \int_0^t H_{\varphi^*}(s) d\Pi_s^*$$

$$(25) \quad \frac{\partial}{\partial\varphi_i} a_{\varphi^*}(t, \cdot) = \int_0^t \left[\frac{\partial}{\partial\varphi_i} \theta(\varphi^*) - H_{\varphi^*}(s) G^{-1} \frac{\partial}{\partial\varphi_i} F(\varphi^*) \right] a_{\varphi^*}(s, \cdot) ds \\ + \int_0^t \left[\theta(\varphi^*) - H_{\varphi^*}(s) G^{-1} F(\varphi^*) \right] \frac{\partial}{\partial\varphi_i} a_{\varphi^*}(s, \cdot) ds + \int_0^t \frac{\partial}{\partial\varphi_i} [H_{\varphi^*}(s)] d\Pi_s^*$$

$$\begin{aligned}
 (26) \quad \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} \alpha_{\varphi^*}(t, \cdot) &= \int_0^t \left[\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} \theta(\varphi^*) - \frac{\partial}{\partial \varphi_j} (H_{\varphi^*}(s)) G^{-1} \frac{\partial}{\partial \varphi_i} F(\varphi^*) \right. \\
 &\quad - \frac{\partial}{\partial \varphi_i} (H_{\varphi^*}(s)) G^{-1} \frac{\partial}{\partial \varphi_j} F(\varphi^*) \\
 &\quad \left. - H_{\varphi^*}(s) G^{-1} \left(\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} F(\varphi^*) \right) \right] \alpha_{\varphi^*}(s, \cdot) ds \\
 &+ \int_0^t \left[\frac{\partial^2}{\partial \varphi_i} \theta(\varphi^*) - \frac{\partial}{\partial \varphi_i} (H_{\varphi^*}(s)) G^{-1} F(\varphi^*) - H_{\varphi^*}(s) G^{-1} \frac{\partial}{\partial \varphi_i} F(\varphi^*) \right] \frac{\partial}{\partial \varphi_j} \alpha_{\varphi^*}(s, \cdot) ds \\
 &+ \int_0^t \left[\frac{\partial^2}{\partial \varphi_j} \theta(\varphi^*) - \frac{\partial}{\partial \varphi_j} (H_{\varphi^*}(s)) G^{-1} F(\varphi^*) - H_{\varphi^*}(s) G^{-1} \frac{\partial}{\partial \varphi_j} F(\varphi^*) \right] \frac{\partial}{\partial \varphi_i} \alpha_{\varphi^*}(s, \cdot) ds \\
 &+ \int_0^t \left[\theta(\varphi^*) - H_{\varphi^*}(s) G^{-1} F(\varphi^*) \right] \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} \alpha_{\varphi^*}(s, \cdot) ds + \int_0^t \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} [H_{\varphi^*}(s)] d\Pi_s^*
 \end{aligned}$$

où $H_{\varphi^*}(s) = Q_{\varphi^*}(s) \cdot {}^t F(\varphi^*) \cdot {}^t G^{-1}$

et $\Pi_t^* = G^{-1} (\Pi_t - \int_0^t F(\varphi^*) \alpha_{\varphi^*}(s, \cdot) ds)$; $t \geq 0$, est un mouvement brownien.

On considère alors l'espace probabilisé :

$$(\mathbb{R}^n \times (\mathbb{R}^n)^d \times (\mathbb{R}^n)^{\frac{d(d+1)}{2}} \times \mathbb{C} , \mathcal{B}_{\mathbb{R}^n} \otimes \mathcal{B}_{(\mathbb{R}^n)^{\frac{d(d+1)}{2}}} \otimes \mathcal{S} , \tilde{\mu}_{\varphi^*})$$

où la mesure μ_{φ^*} est définie par :

$$\tilde{\mu}_{\varphi^*} = N(0, K_*) \otimes \mu_{\varphi^*} ,$$

$N(0, K_*)$ désignant la loi de probabilité gaussienne sur $\mathbb{R}^n \times (\mathbb{R}^n)^d \times (\mathbb{R}^n)^{\frac{d(d+1)}{2}}$ centrée, de matrice de covariances K_* définie par :

$$(27) \quad B_*^\infty \cdot K_* + K_* \cdot {}^t B_*^\infty = -C_*^\infty \cdot {}^t C_*^\infty$$

où, si $B_{\varphi^*}(t)$ et $C_{\varphi^*}(t)$ sont les matrices $n \frac{(d+1)(d+2)}{2} \times n \frac{(d+1)(d+2)}{2}$ et $n \frac{(d+1)(d+2)}{2} \times p$ permettant d'écrire de manière naturelle le système des équations (24), (25) $i=1, d$, (26) $i, j=1, d$ sous la forme vectorielle dans $(\mathbb{R}^n)^{\frac{(d+1)(d+2)}{2}}$:

$$(28) \quad A_{\varphi^*}(t, \cdot) = \int_0^t B_{\varphi^*}(s) A_{\varphi^*}(s, \cdot) ds + \int_0^t C_{\varphi^*}(s) d\tilde{\Pi}_s^*$$

on a posé $B_{\varphi^*}^\infty = \lim_{t \rightarrow +\infty} B_{\varphi^*}(t)$ et $C_{\varphi^*}^\infty = \lim_{t \rightarrow +\infty} C_{\varphi^*}(t)$. En d'autres termes les coefficients des matrices $B_{\varphi^*}^\infty$ et $C_{\varphi^*}^\infty$ s'obtiennent à partir des coefficients correspondants de $B_{\varphi^*}(t)$ et $C_{\varphi^*}(t)$ en y remplaçant $Q_{\varphi^*}(t)$, $\frac{\partial}{\partial \varphi_i} Q_{\varphi^*}(t)$, $\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} Q_{\varphi^*}(t)$ respectivement par $Q_{\varphi^*}^\infty$, $\frac{\partial}{\partial \varphi_i} Q_{\varphi^*}^\infty$ et $\frac{\partial^2}{\partial \varphi_i \partial \varphi_j} Q_{\varphi^*}^\infty$.

Remarquons que les blocs diagonaux $n \times n$ de la matrice $B_{\varphi^*}^\infty$ sont de la forme :

$$\theta(\varphi^*) \quad \text{ou} \quad \theta(\varphi^*) - Q_{\varphi^*}^\infty \cdot {}^t F(\varphi^*) (G \cdot {}^t G)^{-1} F(\varphi^*)$$

et, d'après les hypothèses (L1) et (L2), sont des matrices stables et que de plus les coefficients supérieurs à ces blocs sont tous nuls. Alors la matrice $B_{\varphi^*}^\infty$ est elle-même une matrice stable et l'équation (27) admet bien une solution donnée par :

$$K_{\varphi^*} = \int_0^{+\infty} e^{t B_{\varphi^*}^\infty s} C_{\varphi^*}^\infty \cdot {}^t C_{\varphi^*}^\infty e^{B_{\varphi^*}^\infty s} ds .$$

On définit maintenant sur l'espace probabilisé décrit plus haut le processus $(\tilde{A}_{\varphi^*}(t, \cdot) ; t \geq 0)$ par :

$$\tilde{A}_{\varphi^*}(t, a, x) = A_{\varphi^*}(t, x) \quad \text{pour} \quad (a, x) \in (\mathbb{R}^p \times (\mathbb{R}^p)^d \times (\mathbb{R}^p)^{\frac{d(d+1)}{2}}) \times C, \quad t \geq 0 .$$

Ce processus vérifie encore, relativement à $\tilde{\mu}_{\varphi^*}$:

$$(29) \quad \tilde{A}_{\varphi^*}(t, \cdot) = \int_0^t B_{\varphi^*}(s) \tilde{A}_{\varphi^*}(s, \cdot) ds + \int_0^t C_{\varphi^*}(s) d\tilde{\Pi}_s^*$$

où $\tilde{\Pi}_t^*(a, x) = \Pi_t^*(x)$, $t \geq 0$, est un mouvement brownien.

On définit aussi sur le même espace probabilisé les processus $(\hat{A}_{\varphi^*}(t, \cdot) ; t \geq 0)$ et $(\check{A}_{\varphi^*}(t, \cdot) ; t \geq 0)$ respectivement par :

$$(30) \quad \hat{A}_{\varphi^*}(t, \cdot) = \int_0^t B_{\varphi^*}^\infty \hat{A}_{\varphi^*}(s, \cdot) ds + \int_0^t C_{\varphi^*}^\infty d\tilde{\Pi}_s^*$$

et

$$(31) \quad \check{A}_{\varphi^*}(t, \cdot) = A_{\varphi^*}(0, \cdot) + \int_0^t B_{\varphi^*}^{\infty} \hat{A}_{\varphi^*}(s, \cdot) ds + \int_0^t C_{\varphi^*}^{\infty} d\tilde{\Pi}_s^*$$

où $\check{A}_{\varphi^*}(0, a, x) = a$ et donc $\check{A}_{\varphi^*}(0, \cdot)$ suit, relativement à $\tilde{\mu}_{\varphi^*}$, la loi $N(0, K_{\varphi^*})$. Alors le processus $(\check{A}_{\varphi^*}(t, \cdot); t \geq 0)$ est, relativement à $\tilde{\mu}_{\varphi^*}$, un processus gaussien centré, stationnaire de fonction de covariance définie par :

$$E_{\tilde{\mu}_{\varphi^*}} [\check{A}_{\varphi^*}(t, \cdot) \cdot {}^t\check{A}_{\varphi^*}(s, \cdot)] = e^{(t-s)B_{\varphi^*}^{\infty}} K_{\varphi^*} \quad \text{si } t \geq s.$$

Du lemme II. 1.3.3. on déduit immédiatement la propriété d'approximation suivante :

2.3.3. LEMME : "Sous les hypothèses (L1) et (L2) on a :

$$\lim_{T \rightarrow +\infty} T^{-1/2} \left[\int_0^T d\tilde{\Pi}_s^* \cdot {}^t\hat{A}_{\varphi^*}(s, \cdot) - \int_0^T d\tilde{\Pi}_s^* \cdot {}^t\check{A}_{\varphi^*}(s, \cdot) \right] = 0$$

au sens de la convergence en moyenne quadratique relativement à $\tilde{\mu}_{\varphi^*}$ où $(\hat{A}_{\varphi^*}(t, \cdot); t \geq 0)$ et $(\check{A}_{\varphi^*}(t, \cdot); t \geq 0)$ sont définis respectivement par (30) et (31)."

On s'intéresse maintenant à l'approximation de l'intégrale stochastique $\int_0^T d\tilde{\Pi}_s^* \cdot {}^t\tilde{A}_{\varphi^*}(s, \cdot)$ par $\int_0^T d\tilde{\Pi}_s^* \cdot {}^t\hat{A}_{\varphi^*}(s, \cdot)$:

2.3.4. LEMME : "Sous les hypothèses (L1) et (L2) on a :

$$\lim_{T \rightarrow +\infty} T^{-1/2} \left[\int_0^T d\tilde{\Pi}_s^* \cdot {}^t\tilde{A}_{\varphi^*}(s, \cdot) - \int_0^T d\tilde{\Pi}_s^* \cdot {}^t\hat{A}_{\varphi^*}(s, \cdot) \right] = 0$$

au sens de la convergence en moyenne quadratique relativement à $\tilde{\mu}_{\varphi^*}$ où $(\tilde{A}_{\varphi^*}(t, \cdot); t \geq 0)$ et $(\hat{A}_{\varphi^*}(t, \cdot); t \geq 0)$ sont définis respectivement par (29) et (30).

Démonstration : On écrit $\tilde{R}_{\varphi^*}(t)$ (resp $\hat{R}_{\varphi^*}(t)$) désignant la résolvante de $B_{\varphi^*}(t)$ (resp $B_{\varphi^*}^{\infty}$) :

$$\tilde{A}_{\varphi^*}(t, \cdot) = \tilde{R}_{\varphi^*}(t) \int_0^{t-1} \tilde{R}_{\varphi^*}^{-1}(s) C_{\varphi^*}(s) d\tilde{\Pi}_s^*$$

et

$$\hat{A}_{\varphi^*}(t, \cdot) = \hat{R}_{\varphi^*}(t) \int_0^{t-1} \hat{R}_{\varphi^*}^{-1}(s) C_{\varphi^*}^{\infty} d\tilde{\Pi}_s^*$$

puis
$$\begin{aligned} \tilde{A}_{\varphi^*}(t, \cdot) - \hat{A}_{\varphi^*}(t, \cdot) &= \int_0^t \hat{R}_{\varphi^*}(t) \hat{R}_{\varphi^*}^{-1}(s) (C_{\varphi^*}(s) - C_{\varphi^*}^{\infty}) d\tilde{\Pi}_s^* \\ &+ \int_0^t (\tilde{R}_{\varphi^*}(t) \tilde{R}_{\varphi^*}^{-1}(s) - \hat{R}_{\varphi^*}(t) \hat{R}_{\varphi^*}^{-1}(s)) C_{\varphi^*}(s) d\tilde{\Pi}_s^* . \end{aligned}$$

On déduit facilement du lemme 2.3.2. que :

$$\|B_{\varphi^*}^{\infty} - B_{\varphi^*}(t)\| \leq Ke^{-\alpha(\varphi^*)t}$$

et
$$\|C_{\varphi^*}^{\infty} - C_{\varphi^*}(t)\| \leq Ke^{-\alpha(\varphi^*)t} ,$$

puis que
$$\|\tilde{R}_{\varphi^*}(t) \tilde{R}_{\varphi^*}^{-1}(s) - \hat{R}_{\varphi^*}(t) \hat{R}_{\varphi^*}^{-1}(s)\| \leq Ke^{-\alpha(\varphi^*)(t-s)} .$$

Alors :

$$E_{\tilde{\mu}_{\varphi^*}} \|\tilde{A}_{\varphi^*}(t, \cdot) - \hat{A}_{\varphi^*}(t, \cdot)\|^2 \leq K_1 te^{-2\alpha(\varphi^*)t}$$

et
$$E_{\tilde{\mu}_{\varphi^*}} \left\| \int_0^T d\tilde{\Pi}_s^* \cdot \tilde{A}_{\varphi^*}(s, \cdot) - \int_0^T d\tilde{\Pi}_s^* \cdot \hat{A}_{\varphi^*}(s, \cdot) \right\|^2 \leq K_2$$

puis
$$\lim_{T \rightarrow +\infty} E_{\tilde{\mu}_{\varphi^*}} \|T^{-1/2} \left[\int_0^T d\tilde{\Pi}_s^* \cdot \tilde{A}_{\varphi^*}(s, \cdot) - \int_0^T d\tilde{\Pi}_s^* \cdot \hat{A}_{\varphi^*}(s, \cdot) \right]\|^2 = 0$$

ce qui assure l'assertion du lemme. ■

On est alors en mesure de démontrer le résultat suivant :

2.3.5. LEMME : "Sous les hypothèses (L1) et (L2) on a :

(i)
$$\mu_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T d\tilde{\Pi}_s^* \cdot \tilde{A}_{\varphi^*}(s, \cdot) = 0$$

et (ii)
$$\mathfrak{L}_{\mu_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2} \int_0^T d\tilde{\Pi}_s^* \cdot \tilde{A}_{\varphi^*}(s, \cdot) = M^*$$

où M^* désigne une matrice aléatoire de loi gaussienne $N(0, R^*)$ avec :

$$E(M_{i,j}^* \cdot M_{k,\ell}^*) = R_{i,j,k,\ell}^* = \delta_{i,k} \cdot (K^*)_{j,\ell} ; \quad i, k = \overline{1, p} ; \quad j, \ell = \overline{1, n \frac{(d+1)(d+2)}{2}} ,$$

K^* étant la matrice définie par l'équation (27)."

Démonstration : D'après le lemme II, 1.3.4. on a :

$$\tilde{\mu}_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T d\tilde{\Pi}_s^* \cdot \tilde{A}_{\varphi^*}(s, \cdot) = 0$$

Alors les lemmes 2.3.3 et 2.3.4 assurent que :

$$\tilde{\mu}_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T d\tilde{\Pi}_S^* \cdot {}^t \tilde{A}_{\varphi^*}(s, \cdot) = 0$$

puis

$$\mu_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T d\Pi_S^* \cdot {}^t A_{\varphi^*}(s, \cdot) = 0 .$$

L'assertion (i) du lemme est donc acquise.

De même, d'après la proposition I. 3.2.2 on a :

$$\tilde{\mathfrak{L}}_{\tilde{\mu}_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2} \int_0^T d\tilde{\Pi}_S^* \cdot {}^t \tilde{A}_{\varphi^*}(s, \cdot) = M^*$$

où M^* suit la loi $N(0, R^*)$ définie dans l'énoncé.

Alors les lemmes 2.3.3 et 2.3.4 assurent que :

$$\tilde{\mathfrak{L}}_{\tilde{\mu}_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2} \int_0^T d\tilde{\Pi}_S^* \cdot {}^t \tilde{A}_{\varphi^*}(s, \cdot) = M^*$$

puis

$$\mathfrak{L}_{\mu_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2} \int_0^T d\Pi_S^* \cdot {}^t A_{\varphi^*}(s, \cdot) = M^*$$

ce qui achève la démonstration du lemme. ■

On définit alors la matrice $n(d+1) \times n(d+1)$ S^* comme étant le bloc de la matrice K_* correspondant aux indices (i, j) tels que $i = \overline{1, n(d+1)}$ et $j = \overline{1, n(d+1)}$. En d'autres termes nous pouvons définir la matrice S^* comme étant la matrice $n(d+1) \times n(d+1)$ solution de l'équation :

$$(32) \quad U^* \cdot S^* + S^* \cdot {}^t U^* = -V^* \cdot {}^t V^*$$

où U^* (resp ${}^t V^*$) est le bloc de la matrice B_*^∞ (resp C_*^∞) correspondant aux indices (i, j) tels que $i = \overline{1, n(d+1)}$ et $j = \overline{1, n(d+1)}$ (resp $j = \overline{1, p}$), soit :

- V^* est la matrice $n(d+1) \times p$ définie par :

$$V_{i,j}^* = \begin{cases} (Q_{\varphi^*}^\infty \cdot {}^t F(\varphi^*) \cdot {}^t G^{-1})_{i,j} & \text{si } i = \overline{1, n} ; j = \overline{1, p} \\ \frac{\partial}{\partial \varphi_k} (Q_{\varphi^*}^\infty \cdot {}^t F(\varphi^*) \cdot {}^t G^{-1})_{i-kn, j} & \text{si } i = \overline{kn+1, (k+1)n} ; k = \overline{1,} \\ & j = \overline{1, p} \end{cases}$$

- et U^* est la matrice stable $n(d+1) \times n(d+1)$ définie par :

$$U_{i,j}^* = \begin{cases} (\theta(\varphi^*))_{i,j} & \text{si } i,j = \overline{1,n} \\ (\theta(\varphi^*) - Q_{\varphi^*}^\infty \cdot {}^t F(\varphi^*) (G \cdot {}^t G)^{-1} F(\varphi^*))_{i-kn, j-kn} & \text{si } i,j = \overline{kn+1, (k+1)n}, \quad k = \overline{1,d} \\ \left(\frac{\partial}{\partial \varphi_k} \theta(\varphi^*) - Q_{\varphi^*}^\infty \cdot {}^t F(\varphi^*) (G \cdot {}^t G)^{-1} \frac{\partial}{\partial \varphi_k} F(\varphi^*) \right)_{i-kn, j-(k-1)n} & \text{si } \begin{matrix} i = \overline{kn+1, (k+1)n} \\ j = \overline{(k-1)n+1, kn} \end{matrix}, \quad k = \overline{1,d} \\ 0 & \text{sinon.} \end{cases}$$

On définit aussi successivement les matrices I, D, J, \hat{S} et L , $2dn \times (d+1)n$, $2dp \times 2dn$, $dp \times 2dp$, $pd \times pd$ et $d \times pd$ respectivement par :

$$(33) \quad I_{i,j} = \begin{cases} 1 & \text{si } i = j+kn, \quad k = \overline{0, d-1}, \quad j = \overline{1, n} \\ 1 & \text{si } i = j+(d-1)n, \quad j = \overline{n+1, (d+1)n} \\ 0 & \text{sinon} \end{cases}$$

$$(34) \quad D_{i,j} = \begin{cases} \left(G^{-1} \frac{\partial}{\partial \varphi_k} F(\varphi^*) \right)_{i-kp, j-kn} & \text{si } kp+1 \leq i \leq (k+1)p, \quad kn+1 \leq j \leq (k+1)n, \quad k = \overline{0, d-1} \\ \left(G^{-1} F(\varphi^*) \right)_{i-kp, j-kp} & \text{si } kp+1 \leq i \leq (k+1)p, \quad kn+1 \leq j \leq (k+1)n, \quad k = \overline{d, 2d-1} \\ 0 & \text{sinon} \end{cases}$$

$$(35) \quad J_{i,j} = \begin{cases} 1 & \text{si } i = \overline{1, pd} \text{ et } j = i \text{ ou } j = i+pd \\ 0 & \text{sinon} \end{cases}$$

$$(36) \quad \hat{S}_{i,j} = \begin{cases} \sum_{r,s=1}^{n(d+1)} (JDI)_{i,r} \cdot (JDI)_{j,s} \cdot S_{r,s}^* & \text{si } |i-j| = kp, \quad 0 \leq k \leq d-1, \quad i, j = \overline{1, pd} \\ 0 & \text{sinon} \end{cases}$$

$$(37) \quad L_{i,j} = \begin{cases} 1 & \text{si } j = \overline{(i-1)p+1, ip} , \quad i = \overline{1, d} \\ 0 & \text{sinon} \end{cases}$$

On a alors l'énoncé suivant :

2.3.6. PROPOSITION : "Sous les hypothèses (L1) et (L2) on a :

- (i) $\mu_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} T^{-1} L_T^{(1)}(\varphi^*, \cdot) = 0$
- (ii) $\mu_{\varphi^*} \text{-} \lim_{T \rightarrow +\infty} T^{-1} L_T^{(2)}(\varphi^*, \cdot) = -\Gamma^*$
- (iii) $\mathfrak{L}_{\mu_{\varphi^*}} \text{-} \lim_{T \rightarrow +\infty} T^{-1/2} L_T^{(1)}(\varphi^*, \cdot) = N^*$

où N^* désigne un vecteur aléatoire de loi gaussienne centrée et de matrice de covariance Γ^* donnée par :

$$(38) \quad \Gamma^* = L \cdot \hat{S} \cdot {}^t L ,$$

L étant définie par (37) et \hat{S} est définie par (36)."

Démonstration : Il est facile de voir que le vecteur :

$$L_T^{(1)}(\varphi^*, \cdot) = \begin{pmatrix} L_{T,1}^{(1)}(\varphi^*, \cdot) \\ \vdots \\ L_{T,d}^{(1)}(\varphi^*, \cdot) \end{pmatrix}$$

défini par (22), $i = \overline{1, d}$, pour $\varphi = \varphi^*$, peut s'obtenir à partir de la matrice

$$J.D.I \int_0^T E_{\varphi^*}(s, \cdot) \cdot {}^t d\Pi_s^*$$

où $E_{\varphi^*}(s, \cdot)$ désigne le vecteur des $(d+1)p$ premières composantes de $A_{\varphi^*}(s, \cdot)$ i.e. ${}^t(E_{\varphi^*}(s, \cdot)) = ({}^t \mathcal{A}_{\varphi^*}(s, \cdot), \frac{\partial}{\partial \varphi_1} \mathcal{A}_{\varphi^*}(s, \cdot), \dots, \frac{\partial}{\partial \varphi_d} \mathcal{A}_{\varphi^*}(s, \cdot))$, par la transformation $L \circ \Delta$ où Δ est l'application de $\mathfrak{L}(\mathbb{R}^p; \mathbb{R}^{pd})$ dans \mathbb{R}^{pd} qui à toute matrice $pd \times p$ associe le vecteur de $(\mathbb{R}^p)^d$ dont le $k^{\text{ième}}$ bloc de composantes est la diagonale du $k^{\text{ième}}$ bloc $p \times p$ de cette matrice. D'après le lemme 2.3.5 on a :

$$\mu_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \int_0^T E_{\varphi^*}(s, \cdot) \cdot {}^t d\Pi_s^* = 0$$

et

$$\mathfrak{L}_{\mu_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2} \int_0^T E_{\varphi^*}(s, \cdot) \cdot {}^t d\Pi_s^* = N^*$$

où N^* désigne une matrice aléatoire de loi gaussienne centrée de covariance donnée par :

$$\begin{aligned} E(N_{i,j}^* \cdot N_{k,\ell}^*) &= E(M_{j,i}^* \cdot M_{\ell,k}^*) \\ &= \delta_{j,\ell} \cdot (K^*)_{i,k} \\ &= \delta_{j,\ell} \cdot (S^*)_{i,k} \quad \text{pour } i, k = \overline{1, n(d+1)} \\ &\quad \text{et } j, \ell = \overline{1, p} . \end{aligned}$$

Alors on a aussi :

$$\mu_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} L \circ \Delta(J.D.I. \int_0^T E_{\varphi^*}(s, \cdot) \cdot {}^t d\Pi_s^*) = 0$$

soit :

$$\mu_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1/2} L_T(1)(\varphi^*, \cdot) = 0 ,$$

ce qui fournit l'assertion (i) de la proposition, et un calcul élémentaire montre que :

$$\mathfrak{L}_{\mu_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2} J.D.I. \int_0^T E_{\varphi^*}(s, \cdot) \cdot {}^t d\Pi_s^* = J.D.I. N^*$$

où $J.D.I. N^*$ est une matrice de loi gaussienne centrée de covariance donnée par :

$$E((J.D.I. N^*)_{i,j} (J.D.I. N^*)_{k,\ell}) = \begin{cases} 0 & \text{si } j \neq \ell \\ n(d+1) \\ \sum_{r,s=1} (J.D.I.)_{i,r} (J.D.I.)_{k,s} S_{rs}^* & \text{si } j=\ell \end{cases}$$

pour $i, k = \overline{1, pd}$ et $j, \ell = \overline{1, p}$.

On en déduit immédiatement que :

$$\mu_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1/2} \Delta(J.D.I. \int_0^T E_{\varphi^*}(s, \cdot) \cdot {}^t d\Pi_s^*) = M^*$$

où M^* est un vecteur de loi gaussienne centrée de matrice de covariance \hat{S} définie par (36). L'assertion (iii) s'obtient alors en remarquant que :

$$\mathfrak{L} \underset{\mu_{\varphi^*}}{\text{-lim}} \underset{T \rightarrow +\infty}{T^{-1/2}} L_T^{(1)}(\varphi^*, \dots) = LM^* .$$

Il reste à démontrer l'assertion (ii) de la proposition. Remarquons d'abord que, d'après le lemme 2.3.5, on a :

$$\mu_{\varphi^*} \text{-lim}_{T \rightarrow +\infty} T^{-1} \int_0^T A_{\varphi^*}(s, \dots) \cdot {}^t d\Pi_s^* = 0$$

d'où l'on déduit immédiatement que, pour $i, j = \overline{1, d}$, on a :

$$\mu_{\varphi^*} \text{-lim}_{T \rightarrow +\infty} T^{-1} \int_0^T \left\langle \frac{\partial^2}{\partial \varphi_i \partial \varphi_j} (F(\varphi^*) \mathcal{Q}_{\varphi^*}(t)), (G \cdot {}^t G)^{-1} d\Pi_t^* \right\rangle = 0 .$$

D'après (23) écrite pour $\varphi = \varphi^*$, il suffit donc de montrer que, pour $i, j = \overline{1, d}$, on a :

$$(39) \quad \mu_{\varphi^*} \text{-lim}_{T \rightarrow +\infty} T^{-1} \int_0^T \left\langle \frac{\partial}{\partial \varphi_i} (F(\varphi^*) \mathcal{Q}_{\varphi^*}(t)), (G \cdot {}^t G)^{-1} \frac{\partial}{\partial \varphi_j} (F(\varphi^*) \mathcal{Q}_{\varphi^*}(t)) \right\rangle dt = \Gamma_{i,j}^* .$$

On désigne par $(\check{E}_{\varphi^*}(t, \dots); t \geq 0)$ (resp. $(\hat{E}_{\varphi^*}(t, \dots); t \geq 0)$) (resp. $(\tilde{E}_{\varphi^*}(t, \dots); t \geq 0)$) l'analogie pour $(\check{A}_{\varphi^*}(t, \dots); t \geq 0)$ (resp. $(\hat{A}_{\varphi^*}(t, \dots); t \geq 0)$) (resp. $(\tilde{A}_{\varphi^*}(t, \dots); t \geq 0)$) de $(E_{\varphi^*}(t, \dots); t \geq 0)$ pour $(A_{\varphi^*}(t, \dots); t \geq 0)$.

Le processus $(\check{E}_{\varphi^*}(t, \dots); t \geq 0; \tilde{\mu}_{\varphi^*})$ étant ergodique on a :

$$(40) \quad \lim_{T \rightarrow +\infty} T^{-1} \int_0^T \check{E}_{\varphi^*}(s) \cdot {}^t \check{E}_{\varphi^*}(s) ds = S^* , \quad \tilde{\mu}_{\varphi^*} \text{-presque sûrement.}$$

On évalue alors, $|E|$ désignant $\sum_{i,j} |E_{i,j}|$ pour une matrice E ,

$$(41) \quad \begin{aligned} & E_{\tilde{\mu}_{\varphi^*}} \left| \int_0^T \check{E}_{\varphi^*}(s) \cdot {}^t \check{E}_{\varphi^*}(s) ds - \int_0^T \hat{E}_{\varphi^*}(s) \cdot {}^t \hat{E}_{\varphi^*}(s) ds \right| \\ &= E_{\tilde{\mu}_{\varphi^*}} \left| \int_0^T [\check{E}_{\varphi^*}(s) - \hat{E}_{\varphi^*}(s)] \cdot {}^t \check{E}_{\varphi^*}(s) ds + \int_0^T \hat{E}_{\varphi^*}(s) \cdot {}^t [\check{E}_{\varphi^*}(s) - \hat{E}_{\varphi^*}(s)] ds \right| \\ &\leq K \left[\int_0^T E_{\tilde{\mu}_{\varphi^*}} [\|\check{E}_{\varphi^*}(s) - \hat{E}_{\varphi^*}(s)\|^2] ds \right]^{1/2} \left[\int_0^T E_{\tilde{\mu}_{\varphi^*}} \|\check{E}_{\varphi^*}(s)\|^2 ds \right]^{1/2} + \left[\int_0^T E_{\tilde{\mu}_{\varphi^*}} \|\hat{E}_{\varphi^*}(s)\|^2 ds \right]^{1/2} \end{aligned}$$

$$\text{où } \check{E}_{\varphi^*}(s) - \hat{E}_{\varphi^*}(s) = e^{sU^*} \check{E}_{\varphi^*}(0) .$$

U^* étant une matrice stable, on a :

$$\begin{aligned} E_{\tilde{\mu}_{\varphi^*}} \|\check{E}_{\varphi^*}(s) - \hat{E}_{\varphi^*}(s)\|^2 &= E_{\tilde{\mu}_{\varphi^*}} \|e^{sU^*} \check{E}_{\varphi^*}(0)\|^2 \\ &\leq K e^{-\alpha s} \quad \text{où } \alpha > 0 . \end{aligned}$$

De plus, $(\check{E}_{\varphi^*}(t) ; t \geq 0 ; \tilde{\mu}_{\varphi^*})$ étant un processus stationnaire, on a :

$$E_{\tilde{\mu}_{\varphi^*}} \|\check{E}_{\varphi^*}(t)\|^2 = \text{constante}$$

puis

$$E_{\tilde{\mu}_{\varphi^*}} \|\hat{E}_{\varphi^*}(t)\|^2 \leq K .$$

Alors, il vient, pour le premier membre de l'inégalité (41), la majoration :

$$(K^2 \int_0^{+\infty} e^{-\alpha s} ds) \cdot T^{1/2} ,$$

d'où

$$(42) \quad \tilde{\mu}_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \left[\int_0^T \check{E}_{\varphi^*}(s) \cdot {}^t\check{E}_{\varphi^*}(s) ds - \int_0^T \hat{E}_{\varphi^*}(s) \cdot {}^t\hat{E}_{\varphi^*}(s) ds \right] = 0$$

De même on évalue :

$$\begin{aligned} &E_{\tilde{\mu}_{\varphi^*}} \left| \int_0^T \hat{E}_{\varphi^*}(s) \cdot {}^t\hat{E}_{\varphi^*}(s) ds - \int_0^T \check{E}_{\varphi^*}(s) \cdot {}^t\check{E}_{\varphi^*}(s) ds \right| \\ &\leq K \left[\int_0^T E_{\tilde{\mu}_{\varphi^*}} [\|\hat{E}_{\varphi^*}(s) - \check{E}_{\varphi^*}(s)\|^2] ds \right]^{1/2} \left[\int_0^T E_{\tilde{\mu}_{\varphi^*}} \|\hat{E}_{\varphi^*}(s)\|^2 ds \right]^{1/2} + \left(\int_0^T E_{\tilde{\mu}_{\varphi^*}} \|\check{E}_{\varphi^*}(s)\|^2 ds \right)^{1/2} \end{aligned}$$

qui, d'après la démonstration du lemme 2.3.4, est majoré par :

$$(K' \int_0^{+\infty} t e^{-2\alpha(\varphi^*) \cdot t} dt) T^{1/2} .$$

On a donc aussi :

$$(43) \quad \tilde{\mu}_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \left[\int_0^T \hat{E}_{\varphi^*}(s) \cdot {}^t\hat{E}_{\varphi^*}(s) ds - \int_0^T \check{E}_{\varphi^*}(s) \cdot {}^t\check{E}_{\varphi^*}(s) ds \right] = 0 .$$

Alors (40), (42) et (43) assurent que :

$$\tilde{\mu}_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \left[\int_0^T \check{E}_{\varphi^*}(s) \cdot {}^t\check{E}_{\varphi^*}(s) ds \right] = S^*$$

puis

$$(44) \quad \mu_{\varphi^*} - \lim_{T \rightarrow +\infty} T^{-1} \left[\int_0^T E_{\varphi^*}(s) \cdot {}^tE_{\varphi^*}(s) ds \right] = S^* .$$

Le vecteur de $(\mathbb{R}^p)^d$:

$$\begin{pmatrix} G^{-1} \frac{\partial}{\partial \varphi_1} (F(\varphi^*) \mathcal{O}_{\varphi^*}(t)) \\ \vdots \\ G^{-1} \frac{\partial}{\partial \varphi_d} (F(\varphi^*) \mathcal{O}_{\varphi^*}(t)) \end{pmatrix}$$

peut s'écrire $J.D.I.E_{\varphi^*}(t)$; on a, d'après (44),

$$\mu_{\varphi^*}^{-1} \lim_{T \rightarrow +\infty} T^{-1} \left[\int_0^T JDI E_{\varphi^*}(t) \cdot {}^t JDI E_{\varphi^*}(t) dt \right] = (JDI) S^* \cdot {}^t (JDI)$$

donc :

$$\mu_{\varphi^*}^{-1} \lim_{T \rightarrow +\infty} T^{-1} \left[\int_0^T G^{-1} \frac{\partial}{\partial \varphi_i} (F(\varphi^*) \mathcal{O}_{\varphi^*}(t)) \cdot {}^t \left(G^{-1} \frac{\partial}{\partial \varphi_j} (F(\varphi^*) \mathcal{O}_{\varphi^*}(t)) \right) dt \right] = R_{i,j}$$

où $R_{i,j}$ est le bloc de la matrice $(JDI) S^* \cdot {}^t (JDI)$ correspondant aux indices de lignes $r = \overline{(i-1)p+1, ip}$ et aux indices de colonnes $s = \overline{(j-1)p+1, jp}$, ce pour $i, j = \overline{1, d}$.

Enfin, il vient :

$$\mu_{\varphi^*}^{-1} \lim_{T \rightarrow +\infty} T^{-1} \left[\int_0^T \left(G^{-1} \frac{\partial}{\partial \varphi_i} (F(\varphi^*) \mathcal{O}_{\varphi^*}(t)), G^{-1} \frac{\partial}{\partial \varphi_j} (F(\varphi^*) \mathcal{O}_{\varphi^*}(t)) \right) dt \right] = r_{i,j}$$

où $r_{ij} = \text{Tr} R_{ij}$.

Un calcul élémentaire montre que $r_{ij} = \Gamma_{i,j}^*$, ce qui achève la démonstration compte-tenu de la condition suffisante (39). ■

2.3.7. Remarques :

a) D'après la démonstration précédente la matrice Γ^* définie par (38) peut aussi être définie par :

$$(45) \quad \begin{cases} \Gamma_{i,j}^* = \text{Tr} R_{ij} & ; \quad i, j = \overline{1, d} \\ (R_{ij})_{k,\ell} = [(JDI) S^* \cdot {}^t (JDI)]_{(i-1)p+k, (j-1)p+\ell} & ; \quad k, \ell = \overline{1, p} \\ \text{où } S^*, I, D \text{ et } J \text{ sont respectivement données par (32), (33), (34) et (35)} \end{cases}$$

b) Remarquons que si φ est un système donné de composantes des matrices $\theta(\varphi), \Lambda(\varphi), F(\varphi)$, les matrices dérivées $\frac{\partial}{\partial \varphi_k} \theta(\varphi)$, $\frac{\partial}{\partial \varphi_k} \Lambda(\varphi)$ et $\frac{\partial}{\partial \varphi_k} F(\varphi)$, $k = \overline{1, d}$, sont des matrices dont au plus un coefficient est différent de 0, ce coefficient valant alors 1. Du point de vue du calcul de la matrice Γ^* , le problème est celui de la détermination de la solution S^* de l'équation (32) qui demande le calcul préalable de la matrice $Q_{\varphi^*}^\infty$. Nous reviendrons plus loin sur ces questions.

Nous pouvons maintenant donner un résultat sur l'identifiabilité de φ^* :

2.3.8. PROPOSITION : "Sous les hypothèses (L1) et (L2), si la matrice Γ^* définie par (45) est non singulière, la valeur φ^* est asymptotiquement identifiable en probabilité par une famille d'estimateurs $\{\hat{\varphi}_T; T > 0\}$ vérifiant :

- (i) pour tout $T > 0$, $\hat{\varphi}_T$ est S_T -mesurable,
- (ii) $\lim_{T \rightarrow +\infty} \mu_{\varphi^*} [L_T^{(1)}(\hat{\varphi}_T, \dots) = 0] = 1$, où $L_T^{(1)}(\varphi, \dots)$ est définie par (22),
- (iii) $\mu_{\varphi^*} - \lim_{T \rightarrow +\infty} \hat{\varphi}_T = \varphi^*$
- (iv) $\mathcal{L}_{\mu_{\varphi^*}} - \lim_{T \rightarrow +\infty} T^{-1/2}(\hat{\varphi}_T - \varphi^*) = N(0, \Gamma^{*-1})$.

De plus cette famille d'estimateurs est "unique" au sens que si $\{\tilde{\varphi}_T; T > 0\}$ est une autre famille vérifiant (i), (ii) et (iii) on a aussi :

- (v) $\lim_{T \rightarrow +\infty} \mu_{\varphi^*} [\hat{\varphi}_T = \tilde{\varphi}_T] = 1$."

Démonstration : D'après la proposition 2.3.6. et l'hypothèse d'inversibilité de Γ^* , les hypothèses (E1), (E2) et (E3) de la proposition I, 2.2.1. sont vérifiées. Les assertions (i) jusqu'à (v) seront donc démontrées si nous montrons que l'hypothèse (E4) est satisfaite. La propriété de majoration de l'expression

$$\frac{1}{T} |L_T^{(2)}(\varphi^*, \dots) - L_T^{(2)}(\varphi, \dots)|$$

sera acquise si nous montrons qu'elle est satisfaite par les expressions :

$$(46) \quad \frac{1}{T} \left| \int_0^T A_{\varphi^*}(s) \cdot {}^t[A_{\varphi^*}(s) - A_{\varphi}(s)] ds \right|$$

$$(47) \quad \frac{1}{T} \left| \int_0^T A_{\varphi}(s) \cdot {}^t[A_{\varphi^*} - A_{\varphi}(s)] ds \right|$$

et

$$(48) \quad \frac{1}{T} \left| \int_0^T [A_{\varphi^*}(s) - A_{\varphi}(s)] \cdot {}^t d\Pi_s^* \right|$$

où
$$A_{\varphi^*}(t) = \int_0^t B_{\varphi^*}(s) A_{\varphi^*}(s) ds + \int_0^t C_{\varphi^*}(s) d\Pi_s^*$$

et
$$A_{\varphi}(t) = \int_0^t B_{\varphi}(s) A_{\varphi}(s) ds + \int_0^t C_{\varphi}(s) G^{-1}(F(\varphi^*) \alpha_{\varphi^*}(s) - F(\varphi) \alpha_{\varphi}(s)) ds + \int_0^t C_{\varphi}(s) d\Pi_s^* .$$

On considère alors la différence :

$$(49) \quad \begin{aligned} \alpha_{\varphi}(t) - \alpha_{\varphi^*}(t) &= \int_0^t R_{\varphi}(t) R_{\varphi}^{-1}(s) [(\theta(\varphi) - \theta(\varphi^*)) + H_{\varphi}(s) G^{-1}(F(\varphi^*) - F(\varphi))] \alpha_{\varphi^*}(s) ds \\ &+ \int_0^t R_{\varphi}(t) R_{\varphi}^{-1}(s) [H_{\varphi}(s) - H_{\varphi^*}(s)] d\Pi_s^* \end{aligned}$$

où
$$\alpha_{\varphi^*}(t) = \int_0^t e^{\theta(\varphi^*)(t-s)} H_{\varphi^*}(s) d\Pi_s^*$$

et, pour t supérieur à un certain t_0 , $Q_{\varphi^*}(t)$ étant définie positive (voir [14] théorème 5.1.), $\text{rang}(H_{\varphi^*}(t)) = \text{rang}(F(\varphi^*)) = p$ et donc il existe $H_{\varphi^*}^{-1}(t)$ telle que $H_{\varphi^*}^{-1}(t) \cdot H_{\varphi^*}(t) = \mathbb{1}_p$.

Alors,

$$(50) \quad H_{\varphi^*}^{-1}(t) \alpha_{\varphi^*}(t) - H_{\varphi^*}^{-1}(t_0) \alpha_{\varphi^*}(t_0) = \int_{t_0}^t d\Pi_s^* + \int_{t_0}^t \left[\frac{d}{ds} H_{\varphi^*}^{-1}(s) + H_{\varphi^*}^{-1}(s) \theta(\varphi^*) \right] \alpha_{\varphi^*}(s) ds$$

et si (Φ_t) est un processus à variation bornée tel que l'intégrale stochastique $\int_{t_0}^t \Phi_s d\Pi_s^*$ est définie, on a :

$$\begin{aligned} \int_{t_0}^t \Phi_s d\Pi_s^* &= \int_{t_0}^t \Phi_s d[H_{\varphi^*}^{-1}(s) \alpha_{\varphi^*}(s)] - \int_{t_0}^t \Phi_s \left[\frac{d}{ds} H_{\varphi^*}^{-1}(s) + H_{\varphi^*}^{-1}(s) \theta(\varphi^*) \right] \alpha_{\varphi^*}(s) ds \\ &= \Phi_t H_{\varphi^*}^{-1}(t) \alpha_{\varphi^*}(t) - \Phi_{t_0} H_{\varphi^*}^{-1}(t_0) \alpha_{\varphi^*}(t_0) - \int_{t_0}^t d\Phi_s \cdot H_{\varphi^*}^{-1}(s) \alpha_{\varphi^*}(s) \\ &\quad - \int_{t_0}^t \Phi_s \left[\frac{d}{ds} H_{\varphi^*}^{-1}(s) + H_{\varphi^*}^{-1}(s) \theta(\varphi^*) \right] \alpha_{\varphi^*}(s) ds . \end{aligned}$$

Appliquant cela à $\Phi_t = R_\varphi^{-1}(t)[H_\varphi(t) - H_{\varphi^*}(t)]$ dans (49), compte-tenu des hypothèses (L1) et (L2) et de la continuité des applications $\varphi \rightsquigarrow \theta(\varphi)$, $\varphi \rightsquigarrow \Lambda(\varphi)$ et $\varphi \rightsquigarrow F(\varphi)$, il vient, pour tout $\eta \leq \eta_0$ où

$\{\varphi : \|\varphi - \varphi^*\| \leq \eta_0\} \subset \mathcal{V}(\varphi^*)$, pour un $\alpha > 0$:

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} |a_\varphi(t) - a_{\varphi^*}(t)| \leq g(\eta) \left[\int_0^t e^{-\alpha(t-s)} |a_{\varphi^*}(s)| ds + |a_{\varphi^*}(t)| \right]$$

où $\lim_{\eta \rightarrow 0} g(\eta) = 0$ et $E_{U_{\varphi^*}}(|a_{\varphi^*}(s)|^2) \leq K$ où K ne dépend pas de s .

Alors :

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} |a_\varphi(t) - a_{\varphi^*}(t)| \leq g(\eta) \rho_t$$

où $\lim_{\eta \rightarrow 0} g(\eta) = 0$ et $E_{U_{\varphi^*}}(\rho_t^2) \leq L$ où L est indépendante de t .

On écrit maintenant, $\tilde{R}_{\varphi^*}(t)$ (resp $\tilde{R}_\varphi(t)$) désignant la résolvante de $B_{\varphi^*}(t)$ (resp $B_\varphi(t)$)

$$(51) \quad A_\varphi(t) - A_{\varphi^*}(t) = \int_0^t [\tilde{R}_\varphi(t) \tilde{R}_\varphi^{-1}(s) C_\varphi(s) - \tilde{R}_{\varphi^*}(t) \tilde{R}_{\varphi^*}^{-1}(s) C_{\varphi^*}(s)] d\Pi_s^* \\ + \int_0^t \tilde{R}_\varphi(t) \tilde{R}_\varphi^{-1}(s) C_\varphi(s) G^{-1}(F(\varphi^*) a_{\varphi^*}(s) - F(\varphi) a_\varphi(s)) ds.$$

De la majoration précédente, on déduit alors :

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} |A_\varphi(t) - A_{\varphi^*}(t)| \leq g_1(\eta) \left[\int_0^t e^{-\alpha(t-s)} (|a_{\varphi^*}(s)| + \rho(s)) ds + (|a_{\varphi^*}(t)| + \rho_t) \right]$$

où $\lim_{\eta \rightarrow 0} g_1(\eta) = 0$,

et donc

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} |A_\varphi(t) - A_{\varphi^*}(t)| \leq g_1(\eta) \rho_1(t)$$

où $(g_1, \rho_1(t))$ vérifie les mêmes conditions que (g, ρ) .

Par suite, on a :

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} \left| \int_0^T A_{\varphi^*}(s) \cdot {}^t[A_{\varphi^*}(s) - A_\varphi(s)] ds \right| \leq g_1(\eta) \int_0^T |A_{\varphi^*}(s)| |\rho_1(s)| ds$$

$$\begin{aligned} \text{où } E \int_0^T |A_{\varphi^*}(s)| |\rho_1(s)| ds &\leq M \left[\int_0^T E_{\mu_{\varphi^*}} |A_{\varphi^*}(s)|^2 ds \right]^{1/2} \left[\int_0^T E_{\mu_{\varphi^*}} |\rho_1(s)|^2 ds \right]^{1/2} \\ &\leq N.T \end{aligned}$$

car $E_{\mu_{\varphi^*}} |A_{\varphi^*}(s)|^2$ est bornée.

En définitive, on a :

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} \frac{1}{T} \left| \int_0^T A_{\varphi^*}(s) \cdot {}^t[A_{\varphi^*}(s) - A_{\varphi}(s)] ds \right| \leq g_1(\eta) \tilde{\rho}_1(T)$$

où $\lim_{\eta \rightarrow 0} g_1(\eta) = 0$ et $\tilde{\rho}_1(T)$ est une variable aléatoire d'espérance majorée indépendamment de T .

Ainsi l'expression (46) possède la propriété de majoration demandée. De même on montre que l'expression (47) la vérifie.

On s'intéresse maintenant à l'expression (48). On peut voir aisément que le développement de cette expression à partir de (51) et (50) fait apparaître des termes qui peuvent se mettre sous la même forme que ceux que nous venons d'étudier et qui vérifient les mêmes propriétés, et le terme :

$$\begin{aligned} &\frac{1}{T} \int_{t_0}^T [C_{\varphi}(s) - C_{\varphi^*}(s)] H_{\varphi^*}^{-1}(s) a_{\varphi^*}(s) \cdot {}^t d[H_{\varphi^*}^{-1}(s) \cdot a_{\varphi^*}(s)] \\ &+ \frac{1}{T} \int_{t_0}^T [C_{\varphi}(s) - C_{\varphi^*}(s)] d[H_{\varphi^*}^{-1}(s) a_{\varphi^*}(s)] \cdot {}^t [H_{\varphi^*}(s) a_{\varphi^*}(s)] . \end{aligned}$$

On l'évalue par :

$$\begin{aligned} &\frac{1}{T} [C_{\varphi}(T) - C_{\varphi^*}(T)] H_{\varphi^*}^{-1}(T) a_{\varphi^*}(T) \cdot {}^t a_{\varphi^*}(T) \cdot {}^t H_{\varphi^*}^{-1}(T) \\ &- \frac{1}{T} [C_{\varphi}(t_0) - C_{\varphi^*}(t_0)] H_{\varphi^*}^{-1}(t_0) a_{\varphi^*}(t_0) \cdot {}^t a_{\varphi^*}(t_0) \cdot {}^t H_{\varphi^*}(t_0) \\ &- \frac{1}{T} \int_{t_0}^T \frac{d}{ds} [C_{\varphi}(s) - C_{\varphi^*}(s)] H_{\varphi^*}^{-1}(s) a_{\varphi^*}(s) \cdot {}^t a_{\varphi^*}(s) \cdot {}^t H_{\varphi^*}^{-1}(s) ds \\ &- \frac{1}{T} \times \int_{t_0}^T [C_{\varphi}(s) - C_{\varphi^*}(s)] ds \end{aligned}$$

et on peut donc le majorer sur $\|\varphi - \varphi^*\| \leq \eta$ par :

$$g_2(\eta) \left\{ \frac{1}{T} |\alpha_{\varphi^*}(T)|^2 + \frac{1}{T} |\alpha_{\varphi^*}(t_0)|^2 + \frac{1}{T} \int_{t_0}^T |\alpha_{\varphi^*}(s)|^2 ds + 1 \right\}$$

où $\lim_{\eta \rightarrow 0} g_2(\eta) = 0$.

Comme $E_{\varphi^*} (|\alpha_{\varphi^*}(s)|^2) \leq K$, on obtient le majorant :

$$g_2(\eta) \times \rho_2(T)$$

où (g_2, ρ_2) vérifie les mêmes conditions que $(g_1, \tilde{\rho}_1)$.

En définitive, on a :

$$\sup_{\|\varphi - \varphi^*\| \leq \eta} \frac{1}{T} |L_T^{(2)}(\varphi^*) - L_T^{(2)}(\varphi)| \leq g_3(\eta) \rho_3(T)$$

où $\lim_{\eta \rightarrow 0} g_3(\eta) = 0$ et $\rho_3(T)$ est d'espérance majorée par une constante ne dépendant pas de T . Ainsi l'hypothèse (E4) de la proposition I, 2.2.1. est vérifiée, ce qui achève la démonstration. ■

2.3.9. Remarques :

Les techniques de démonstration de A.V. BALAKRISHNAN [7] lui permettent seulement d'assurer que la limite, quand T tend vers l'infini, de $\frac{1}{T} L_T^{(2)}(\varphi, \cdot)$ existe en un certain sens de convergence uniforme et que si cette limite pour $\varphi = \varphi^*$ est non singulière alors pour tout voisinage de φ^* et tout x appartenant à un ensemble de probabilité arbitrairement grande on peut trouver un T suffisamment grand tel que l'équation $L_T^{(1)}(\varphi, x) = 0$ ait une solution dans ce voisinage. En d'autres termes cet auteur assure uniquement les conclusions du lemme I, 2.2.3, encore ne se préoccupe-t-il pas de l'existence d'une version régulière de la fonction de vraisemblance ce qui nous semble important compte-tenu de ce que le problème d'estimation est par essence un problème trajectorien. Nous pensons donc avoir amélioré notablement les conclusions du travail [7] :

- d'une part, nous justifions l'utilisation de la méthode du maximum de vraisemblance en donnant une méthode d'estimation préalable des paramètres de covariance du bruit d'observation dans le modèle (12),

- d'autre part, nous donnons une caractérisation algébrique de la matrice Γ^* intervenant dans la condition d'identifiabilité : " Γ^* est non singulière", rendant la vérification de la dite condition possible,

- enfin, nous assurons l'existence d'une famille d'estimateurs convergeant en probabilité et asymptotiquement normalement distribuée.

2.4. Sur la condition d'identifiabilité et le calcul de l'estimateur de φ .

Il est clair qu'en pratique il n'est pas question de vérifier précisément la condition d'identifiabilité puisque la matrice Γ^* concernée met justement en jeu la valeur φ^* inconnue du paramètre φ à estimer. Une situation concrète où l'étude prend véritablement son sens est celle où le domaine Φ des valeurs possibles de φ est un voisinage de φ^* éventuellement parce qu'une estimation a priori est accessible par l'analyse du phénomène modélisé ou parce que la modélisation provient d'une identification préalable des dimensions du système par une méthode du type "réalisation stochastique" (voir par exemple [26] B.L. HO et R.E. KALMAN, [28] E. IRVING, [21] P. FAURRE, [44] J.R. RISSANEN ...) ayant fourni une première estimation des coefficients du modèle.

La condition à demander est alors, de façon naturelle, que dans ce voisinage de φ^* la matrice $\Gamma(\varphi)$, définie par les analogues des équations (32), (33), (34), (35) et (45) obtenues en remplaçant, là où figure φ^* , φ^* par φ , soit une matrice non singulière. La vérification repose évidemment sur le calcul des matrices Q_φ^∞ et $S(\varphi)$ où $S(\varphi)$ est l'analogue pour φ de S^* pour φ^* soit les matrices vérifiant respectivement les équations :

$$(52) \quad \theta(\varphi)Q + Q \cdot {}^t\theta(\varphi) + \Lambda(\varphi) \cdot {}^t\Lambda(\varphi) - Q \cdot {}^tF(\varphi)(G \cdot {}^tG)^{-1}F(\varphi)Q = 0$$

et

$$(53) \quad U(\varphi)S + S \cdot {}^tU(\varphi) = -V(\varphi) \cdot {}^tV(\varphi)$$

où $U(\varphi)$ et $V(\varphi)$ sont les analogues pour φ de U^* et V^* pour φ^* .

Les méthodes constructives de résolution de l'équation (52) sont basées essentiellement sur des algorithmes de type NEWTON-RAPHSON (voir par exemple [54] W.M. WONHAM) :

$$(52)' \quad (\theta(\varphi) - Q_n \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi)) Q_{n+1} + Q_{n+1} \cdot {}^t (\theta(\varphi) - Q_n \cdot {}^t F(\varphi) (G \cdot {}^t G)^{-1} F(\varphi)) \\ + \Lambda(\varphi) \cdot {}^t \Lambda(\varphi) + Q_n \cdot {}^t F(\varphi) \cdot (G \cdot {}^t G)^{-1} F(\varphi) Q_n = 0 .$$

L'équation (53) dont la solution est donnée par :

$$(53)' \quad S(\varphi) = \int_0^{+\infty} e^{t U(\varphi) s} V(\varphi) \cdot {}^t V(\varphi) e^{U(\varphi) s} ds$$

est une équation linéaire.

De même la construction de l'estimateur $\hat{\varphi}_T$ repose sur la mise en oeuvre d'une méthode numérique de résolution de l'équation :

$$(54) \quad L_T^{(1)}(\varphi, x) = 0 .$$

On peut, comme au § 1.3. précédent, envisager un algorithme du type NEWTON-RAPHSON :

$$(54)' \quad \varphi_T^{(n+1)}(x) = \varphi_T^{(n)}(x) - [L_T^{(2)}(\varphi_T^{(n)}(x), x)]^{-1} [L_T^{(1)}(\varphi_T^{(n)}(x), x)]$$

qui demande la détermination des dérivées premières et secondes de la fonction de Log-vraisemblance, basée elle-même sur le calcul de la solution $Q_\varphi(t)$ de (17) (voir encore [54] W.M. WONHAM). Dans [27] K.W. ILIFF donne des résultats numériques d'expériences obtenus par cette approche.

Du point de vue du problème pratique de la résolution de (54) il nous semble encore que nos résultats théoriques sont intéressants : il paraît en effet raisonnable, compte-tenu de ce que $\frac{1}{T} L_T^{(2)}(\varphi, \cdot)$ converge vers $\Gamma(\varphi)$, de substituer à l'algorithme (54)' l'algorithme :

$$(54)'' \quad \varphi_T^{(n+1)}(x) = \varphi_T^{(n)}(x) - [\Gamma(\varphi_T^{(n)}(x))]^{-1} \left[\frac{1}{T} \nabla_T(\varphi_T^{(n)}(x), x) \right]$$

où $\nabla_T(\varphi, x)$ représente l'analogue de $L_T^{(1)}(\varphi, x)$ obtenu en substituant dans les équations la définissant $Q_\varphi(t)$ par Q_φ^∞ . Le calcul des intégrales stochastiques intervenant dans ces équations nécessite lui-aussi l'utilisation de techniques d'approximation : A.V. BALAKRISHNAN étudie ce problème dans [8], ainsi que E. WONG et M. ZAKAI dans [53] et M.F. ALLAIN dans [2a] et [2b].

§ 3. - SUR LA MODELISATION ET L'IDENTIFICATION DES SYSTEMES STOCHASTIQUES A TEMPS CONTINU. QUELQUES REMARQUES POUR CONCLURE.

Le but de ce dernier paragraphe, délibérément non mathématique, est de discuter de quelques aspects du problème de la modélisation et de l'identification des systèmes stochastiques à temps continu et de faire à cette occasion quelques remarques permettant d'envisager certains prolongements à notre travail. Schématiquement, le problème de la construction d'un modèle descriptif d'un phénomène stochastique (à temps continu) à partir de données d'observations comprend deux étapes. La première étape consiste en le choix d'un type de modèles dont il soit raisonnable d'admettre que l'un d'eux représente bien le phénomène observé : cette étape dont nous dirons, pour fixer les idées, qu'elle est celle de la modélisation, repose sur des considérations liées à la nature physique (ou sociologique, ...) du phénomène observé permettant de dégager un cadre d'hypothèses (linéarités ou non linéarités, indépendances ou interactions, continuités ou discontinuités, ...) aboutissant à la construction d'une structure mathématique qui soit une représentation simplifiée du système sur lequel portent les observations et qui rende compte des caractéristiques dominantes de ce système. Une classe de modèles étant fixée, la seconde étape a pour objet la détermination dans cette classe de celui des modèles qui décrit au mieux le phénomène : cette étape, dont nous dirons qu'elle est celle de l'identification, peut mettre en jeu des méthodes statistiques du type estimation et test. Dans [13], G.E.P. BOX et G.M. JENKINS consacrent les chapitres VI et VII respectivement à chacune de ces étapes pour le cas des phénomènes à temps discret. Il est clair qu'en pratique la modélisation et l'identification sont intimement liées, ne serait-ce que parce que l'on ne retiendra comme base de travail une classe de modèles que dans la mesure où on saura y traiter le problème d'estimation de paramètres. Il est clair aussi que le travail que nous venons de présenter s'inscrit tout entier dans le cadre de la phase d'identification : nous avons en effet systématiquement supposé que les observations faites étaient en-

gendrées par un certain type de modèles différentiels stochastiques parce que l'analyse du phénomène concerné et une phase préalable de modélisation avaient conduit à ce choix.

3.1. Sur l'identification des dimensions.

Dans de nombreuses situations, la structure dégagée lors de l'étape de modélisation fait apparaître l'observation dont on dispose comme l'image en dimension p (connue !) de l'état non observable du système de dimension n (inconnue), $n \geq p$: un problème important est alors de déterminer une dimension n (en général la plus petite possible) telle que la modélisation soit satisfaisante.

Dans [14] chapitre V, R.S. BUCY et P.D. JOSEPH, donnent une solution à ce problème pour un modèle du type

$$\begin{cases} dX_t = FX_t dt + GdW_t \\ Y_t = HX_t \end{cases}$$

où (Y_t) est un processus stationnaire, supposant que F est une matrice stable et que les couples (H,F) et (G,F) sont respectivement contrôlable et observable, et utilisant comme critère d'adéquation le fait que la fonction de corrélation dans le modèle coïncide avec celle calculée sur l'observation.

Dans [20], E.D. EYMAN et T. KERR étudient le cas du modèle plus général

$$\begin{cases} dX_t = FX_t dt + MV_t dt + GdW_t \\ Y_t = HX_t + m \end{cases}$$

où (V_t) est une entrée déterministe à la disposition de l'observateur, dans le même cadre d'hypothèses.

Enfin l'algorithme de HO (voir [26], [21], ...) fournit une solution du problème dans le cas du modèle de filtrage linéaire que nous ve-

nous d'étudier, sous des hypothèses de même nature.

Nous pensons que, dans ce dernier cas, en particulier, le problème de choix d'une dimension de l'état non observable peut en fait se formuler sous la forme d'un problème d'estimation ou de test statistique. En effet, la covariance du bruit d'observation ayant été identifiée, la structure statistique est :

$$(C, S_T, \{\mu_{(\theta_n, \Lambda_n, F_n)}^T ; (\theta_n, \Lambda_n, F_n) \in \Delta_n, p \leq n \leq n_0\})$$

où n_0 est une borne supérieure à la dimension n imposée par exemple par des contraintes techniques, et pour tout $n \in \{p, \dots, n_0\}$, Δ_n représente l'ensemble des valeurs possibles du triplet $(\theta_n, \Lambda_n, F_n)$, $\mu_{\theta_n, \Lambda_n, F_n}$ désignant la loi de l'observation pour la valeur $(\theta_n, \Lambda_n, F_n)$. Il est clair que cette structure statistique est encore une structure dominée dans laquelle on peut envisager l'estimation de $(\theta_n, \Lambda_n, F_n)$ par la méthode du maximum de vraisemblance i.e. par la recherche de $(\hat{n}, \hat{\theta}, \hat{\Lambda}, \hat{F})$ tel que $(\hat{\theta}(x), \hat{\Lambda}(x), \hat{F}(x)) \in \Delta_{\hat{n}(x)}$ et :

$$L_T(\hat{\theta}(x), \hat{\Lambda}(x), \hat{F}(x), x) = \sup\{L_T(\theta_n, \Lambda_n, F_n, x) ; p \leq n \leq n_0, (\theta_n, \Lambda_n, F_n) \in \Delta_n\}.$$

On peut d'autre part envisager le choix entre deux valeurs possibles n_1 et n_2 de n par un test de l'hypothèse $n = n_1$ contre l'hypothèse $n = n_2$ ($(n_1, n_2) \in \{p, \dots, n_0\}^2$) dans cette structure statistique, construit de la manière suivante :

$$\text{"Si } L_T(\hat{\theta}_1(x), \hat{\Lambda}_1(x), \hat{F}_1(x), x) \geq L_T(\hat{\theta}_2(x), \hat{\Lambda}_2(x), \hat{F}_2(x), x)$$

on décide que $n = n_1$ ", où pour $i = 1, 2$ $(\hat{\theta}_i, \hat{\Lambda}_i, \hat{F}_i)$ désigne l'estimateur de (θ, Λ, F) dans la structure statistique

$$(C, S_T, \{\mu_{(\theta_{n_i}, \Lambda_{n_i}, F_{n_i})}^T ; (\theta_{n_i}, \Lambda_{n_i}, F_{n_i}) \in \Delta_{n_i}\}).$$

3.2. Sur le choix de l'intégrale stochastique et la nature du bruit.

Dans le présent travail nous n'avons considéré que des modèles engendrant des processus à trajectoires continues où les équations diffé-

rentielles stochastiques mettent en jeu un processus de mouvement brownien, l'intégration stochastique étant prise au sens de K. ITO.

D'une part certains phénomènes à trajectoires continues sont susceptibles de modélisations plus satisfaisantes lorsque l'intégration stochastique est prise au sens de R.L. STRATONOVITCH [48] (par exemple : mouvement sur une orbite circulaire d'une particule soumise à des forces aléatoires pour lequel la modélisation la plus naturelle par une équation différentielle stochastique conduit au sens de ITO, à une diffusion dans tout le plan alors que son interprétation au sens de STRATONOVITCH donne bien comme support du mouvement l'orbite).

D'autre part certains phénomènes à trajectoires discontinues sont modélisables par des équations différentielles stochastiques faisant intervenir des bruits de type poissonien ou plus généralement des martingales discontinues.

Dans ces deux directions, compte-tenu des relations existant entre l'intégration au sens de ITO et celle au sens de STRATONOVITCH et compte-tenu de résultats sur l'absolue continuité des lois de processus régies par des équations différentielles par rapport à des martingales discontinues (voir par exemple [51], J.H. VAN SHUPPEN), il nous semble possible de prolonger notre travail.

3.3. Sur l'identification : estimation, domaines de confiance, tests.

Dans notre travail, nous avons essentiellement abordé le problème d'estimation dans une structure statistique issue d'une étape de modélisation, par la méthode du maximum de vraisemblance. Nous avons toutefois remarqué au II-§ 3.4., à propos d'un cas simple, qu'une approche bayésienne conduit à un problème de filtrage non linéaire. On peut voir encore que dans le cas d'un modèle différentiel général du type

$$dX_t = \alpha(\varphi, t, X)dt + \sigma(\varphi, t, X)dW_t$$

l'approche bayésienne du problème d'estimation de φ , conduit au modèle de filtrage non linéaire

$$\begin{cases} d\varphi_t = 0 \\ dX_t = a(\varphi_t, t, X)dt + \sigma(\varphi_t, t, X)dW_t . \end{cases}$$

L'obtention de résultats statistiques par une telle approche repose donc sur les progrès de la théorie du filtrage non linéaire ([36] R.S. LIPCER et A.N. SHIRYAYEV, ...).

La statistique mathématique classique traite de trois types de problèmes : l'estimation ponctuelle, l'estimation par domaines de confiance et les tests. Nous n'avons étudié que le problème d'estimation : les résultats que nous avons obtenus, en particulier sur la loi asymptotique des estimateurs, permettent d'envisager la construction de domaines de confiances et de tests, ces derniers, dans les cas les plus simples, ayant déjà trait à des structures statistiques exponentielles incomplètes.

BIBLIOGRAPHIE

- [1] J. AITCHINSON - S.D. SILVEY - "Maximum likelihood estimation of parameters subject to constraint". Ann. of Math. Statist., Vol. 29, (1955), p; 813-828.
- [2a] M.F. ALLAIN - "Sur quelques types d'approximation des solutions d'équations différentielles stochastiques". Thèse 3e cycle, Rennes (1974).
- [2b] M.F. ALLAIN - "Approximation, par un processus de diffusion, des oscillations, autour d'une valeur moyenne, d'un processus de Markov de sauts pur". Séminaire de Rennes (1975).
- [3] T.W. ANDERSON - "The statistical analysis of time series". John Wiley, New-York, (1971).
- [4] L. ARNOLD - "Stochastic Differential equations. Theory and applications". John Wiley, New-York, (1974).
- [5] K.J. ASTRÖM - P. EYCKHOFF - "Systems identification. A Survey". Proceedings of the IFAC Symposium on Identification and Parameter Estimation, Prague, (1970), p. 1-38.
- [6] A. BAGCHI - "Continuous Time Systems Identification with unknown noise covariance". Automatica, Vol. 11, (1975), p. 533-535.
- [7] A.V. BALAKRISHNAN - "Stochastic Differential Systems I". Lectures Notes in Economics and Mathematical Systems, Springer-Verlag, Vol. 84, (1973).
- [8] A.V. BALAKRISHNAN - "On the approximation of Ito integrals using band-limited Processes". SIAM J., Control, Vol. 12, n° 2, (1974), p. 237-251.
- [9] J.R. BARRA - "Notions Fondamentales de Statistique Mathématique". Dunod Université, Paris, (1971).
- [10] V.E. BENES - "Existence of optimal stochastic control laws". SIAM J. Control, Vol. 9, n° 3, (1971), p. 446-472.
- [11] P. BILLINGSLEY - "Statistical inference for Markov Processes". University of Chicago Press, (1961).
- [12] P. BILLINGSLEY - "Convergence of probability measures". John Wiley, New-York, (1968).

- [28] E. IRVING - "Identification des Systèmes II". Note E.D.F., n° H.R. 8202 (1968).
- [29] T. KAILATH - "An Innovation Approach to least Square Estimation, Part I : Linear filtering in Additive White Noise". I.E.E.E. Trans. on Automatic Control, Vol. AC 13, (1968), p. 646-655.
- [30] A. LE BRETON - "Estimation des paramètres d'une équation différentielle stochastique vectorielle linéaire". C.R.A.S. Paris, Série A, t. 279, (1974), p. 289-292.
- [31] A. LE BRETON - "Parameter estimation in a linear differential equation". Trans. of 7th Prague Conference and 1974 E.M.S., à paraître.
- [32] A. LE BRETON - "Estimation de paramètres dans un modèle de système dynamique à état et observation régis par des équations différentielles stochastiques". C.R.A.S. Paris, Série A, t. 280, (1975), p. 1377-1380.
- [33] A. LE BRETON - "On continuous and discrete sampling for parameter estimation in diffusion type processes". Proceedings of the International Symposium on stochastic systems, Lexington, U.S.A., 1975, à paraître dans Math. Prog. Studies.
- [34] A. LE BRETON - "Estimation dans un modèle différentiel stochastique à coefficients dépendant différentiablement des paramètres". Séminaire I.R.I.A., (1975).
- [35] E.B. LEE - L. MARKUS - "Foundations of Optimal Control Theory". SIAM Series in Applied Mathematics, John Wiley, New-York, (1967).
- [36] P.S. LIPCHER - A.N. SHIRYAYEV - "Statistics of stochastic processes (non linear filtration and related topics)". Moscou, (1974) (en russe).
- [37] LOPEZ TOLEDO - "Optimal inputs for identification of stochastic systems". M.I.T. Ph. D. (1974).
- [38] A. MEGE-YTHIER - A. MUYSHONDT - "Mise en oeuvre d'une méthode d'estimation de paramètres dans la structure statistique associée à l'observation de certaines fonctions aléatoires. Simulation, aspects théoriques et pratiques". Projet 3e année ENSIMA, Grenoble (1975).
- [39] J. MEMIN - "Sur quelques problèmes fondamentaux de la théorie du filtrage". Thèse, Université de Rennes, (1974).
- [40] J. NEVEU - "Intégrales stochastiques et applications". Cours de 3e cycle, Université de Paris VI (1971-1972).

- [41] A.A. NOVIKOV - "Sequential estimation of the parameters of diffusion processes". *Teor. Verojatn.i.Primenen*, Vol. 16, (1971), p. 394-396. (en russe).
- [42] PHAM DINH TUAN - "Estimation et test dans les modèles paramétriques de processus stationnaires". Thèse, Université de Grenoble, (1975).
- [43] P. PRIOURET - "Processus de Diffusion et Equations Différentielles Stochastiques". Cours d'Eté de St-Flour. Lecture Notes in Math. Vol. 390, Springer Verlag, (1974).
- [44] J.R. RISSANEN - "Recursive identification of linear systems". *SIAM J. CONTROL*, vol. 9, n° 3, (1971), p. 420-430.
- [45] Yu. A. ROZANOV - "Stationary Random Processes". Holden Day, San Francisco, (1967).
- [46] T.L. SAATY - "Modern non linear equations". Mac Graw-Hill, New-York, (1967).
- [47] A.N. SHIRYAYEV - "Statistics of diffusion type processes". Proceedings of 2nd Japan - USSR Symposium on Probability Theory. Lectures Notes, Springer-Verlag, Vol. 330, (1973), p. 397-411.
- [48] R.L. STRATONOVITCH - "Conditional Markov processes and their application to the theory of optimal control". Elsevier, New-York, (1968).
- [49] A.F. TARASKIN - "The asymptotic normality of stochastic vector integrals and estimates of the drift parameters of a multidimensional diffusion process". *Teor. Verojatn.i. Math. Statist.*, Vol. 2, (1970), p. 205-220. (en russe).
- [50] A.F. TARASKIN - "Some limit theorems for stochastic integrals". *Theory of Stochastic Processes*, J. Wiley, Vol. 1, (1974), p. 136-151. (traduit du russe).
- [51] J.H. VAN SCHUPPEN - "Estimation theory for continuous time processes, a martingale approach". *El. Res. Lab. Mem. N° ERL-M405*, College of Eng., Berkeley, Calif., (1973).
- [52] J. WARGA - "Optimal control of differential and functional equations". Academic Press, New-York, (1972).
- [53] E. WONG - M. ZAKAI - "Riemann-Stieljes approximations of stochastic integrals". *Zeits für Wahrscheinlichkeits theorie*, Vol. 12, n° 2, (1969), p. 87-97.

- [54] W.M. WONHAM - "On a matrix Riccati equation of stochastic control". SIAM J. Control, Vol. 6, n° 4, (1968).
- [55] W.M. WONHAM - "On the separation theorem of stochastic control". SIAM J. Control, Vol. 6, n° 2, (1968), p. 312-326.

-:-:-:-

Dernière page d'une thèse

VU

Grenoble, le

Le Président de la thèse


J.R. BARRA

VU, et permis d'imprimer,

Grenoble, le

Le Président de l'Institut
National Polytechnique

Ph. TRAYNARD
Président
de l'Institut National Polytechnique



Le Président de l'Université
Scientifique et Médicale

