1

of/de

1

National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service

Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

INTERACTING FLEMING-VIOT PROCESSES
AND RELATED MEASURE-VALUED PROCESSES

by

Jean Vaillancourt, B.Sc., M.Sc.

A thesis submitted to
The Faculty of Graduate Studies and Research
in partial fulfilment of
the requirements for the degree of
Doctor of Philosophy

Department of Mathematics and Statistics

Carleton University

Ottawa, Ontario

August, 1987

© copyright

1987, Jean Vaillancourt

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-39360-2

The undersigned hereby recommend to
The Faculty of Graduate Studies and Research
acceptance of the thesis,

"INTERACTING FLEMING-VIOT PROCESSES
AND RELATED MEASURE-VALUED PROCESSES"

submitted by

JEAN VAILLANCOURT, B.Sc., M.Sc.

in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

Chairman, Department of Mathematics
and Statistics

Thesis Supervisor

External Examiner

Carleton University

August, 1987

ABSTRACT

Necessary and sufficient conditions are given for the solution of a large class of martingale problems on $C([0, \infty); (\mathbb{R}^d)^n)$ to be n -exchangeable, in terms of the drift and diffusion coefficients.

Under some mild growth conditions, it is proved that the sequence of empirical processes associated with a triangular array of exchangeable diffusions is tight, thereby proving the existence of at least one solution for a large class of martingale problems on $C([0, \infty); P(\mathbb{R}^d))$.

A class of interacting Fleming-Viot processes is constructed, and then characterized as the unique solution to a martingale problem on $D([0, \infty); P(\mathbb{R}^d)^n)$. For a subclass of such processes, a scaling theorem is proved.

A new condition for tightness in $P(D([0, \infty); P(\mathbb{R}^d)^n))$ is given, in terms of the tightness of a small collection of related integrals in $P(D([0, \infty); \mathbb{R}))$.

ACKNOWLEDGEMENTS

My thanks first go to Dr. Donald Dawson, for showing great patience during the entire course of researching this thesis. I have benefitted also from the friendship of Mr. Whitney Broussard, Dr. Colleen Cutler, Dr. Sue Seager, and Mr. Byron Schmuland, with whom I have had many stimulating conversations. To my parents, I send "un coup-de chapeau" for all their encouragements.

This thesis was typed by Mrs. Gill Murray, who did an excellent job, cheerful and errorfree.

Finally, I would like to thank Janice Jeffs, for providing the right blend of humor and support over the years.

TABLE OF CONTENTS

Acceptance Sheet	ii
Abstract	iii
Acknowledgements	iv
Table of Contents	v
Introduction	1
Chapter 1. Weak Convergence of Probability Measures	3
1. The Space $P(S)$	4
2. The Spaces $C([0, \infty); S)$ and $D([0, \infty); S)$	10
3. D-Semimartingales and the Joffe-Métivier Criterion	16
Chapter 2. Exchangeable Systems of Particles	20
1. Exchangeable Sequences of Processes	20
2. Martingale Problems: Uniqueness and Duality	28
3. Exchangeable Sequences of Diffusions	31
4. On the Existence of Solutions for some Martingale Problems on $C([0, \infty); P(\mathbb{R}^d))$	38
Chapter 3. Interacting Fleming-Viot Processes	53
1. Function-Valued Duals for Interacting Measure-Valued Processes	53
2. Interacting Ohta-Kimura Processes	57
3. Interacting Fleming-Viot Processes: Existence	61
4. Interacting Fleming-Viot Processes: Uniqueness	69
5. The Scaling Theorem	74
Appendix	78
Bibliography	83

INTRODUCTION

The study of measure-valued stochastic processes, preluded by the work of W. Feller on the evolution on the Galton-Watson model in genetics, arises from the need to describe mathematically the global behavior of large collections of interdependent individuals, be they particles or members of a population. (For more information on the subject, the reader should consult [DA2] and [EK]).

This thesis is concerned with the mathematical construction and analysis of two classes of measure-valued processes: those arising from the study of systems of indistinguishable interacting particles (Chapter 2) and a new class of interacting measure-valued processes, originating with the stepwise mutation model of Ohta and Kimura [OK] for estimating the number of electrophoretically detectable alleles in a finite population.

In Chapter 2, after defining the notion of exchangeability for stochastic processes (as done in [AL]), the mathematical equivalent of indistinguishability of motion, we give necessary and sufficient conditions for this property to occur first, in the case of Markov processes; in terms of the transition function, then, in the case of non-explosive diffusions, in terms of the drift and diffusion coefficients (sections 1 and 3, respectively).

In Section 4, we prove that, under some mild conditions on the diffusion coefficients, the sequence of empirical processes associated with a triangular array of exchangeable diffusions is tight, giving rise to at least one solution for certain martingale problems on

$\mathcal{C}([0, \infty) : P(\mathbb{R}^d))$. Such existence theorems are obtained, for instance, in [DA1], [LE], [OE], where the theory of stochastic differential equations used, although yielding weak laws of large numbers and characterizations of the unique limit, requires stronger assumptions for existence than the martingale problem formulation used here.

In Chapter 3, we construct a class of interacting population models (section 2) and prove their weak convergence when the population size increases to infinity, the evolution is speeded up by a factor n^2 , the incremental effects of a change in each population's distribution of characteristics under study is scaled down by a factor n^{-1} and the level of interaction between populations is decreased by a factor n^{-1} (sections 3 and 4). The limits, characterized as unique solutions to certain martingale problems using the method of duality (as in [DH1] and [DK]), are interacting Fleming-Viot processes (see [FV] for the original characterization of the single Fleming-Viot process).

In Section 5, we prove a scaling theorem for these interacting Fleming-Viot processes, using the technique of duality, as was done in the proof of the scaling theorem for a single Fleming-Viot process in [DH2].

The appendix contains a proof of a criterion for tightness in $D([0, \infty) : P(\mathbb{R}^d)^n)$ used in Chapter 3.

Undefined terminology and standard results from basic analysis, probability and stochastic processes theory may be found respectively in [LA], [BI] and [EK].

CHAPTER 1

WEAK CONVERGENCE OF PROBABILITY MEASURES

In this chapter, we review some important well-known results on the weak convergence of probability measures, including a recent general criterion for the tightness of so-called D-semimartingales, due to Joffe and Métivier [JM], which is used in a later chapter in the proof of a weak law of large numbers for sequences of exchangeable diffusions. For a thorough presentation of weak convergence (and much more), the reader may consult the book of Ethier and Kurtz [EK].

We begin by introducing some basic definitions and notation.

Throughout this work,

- $(\Omega, \mathcal{F}, \mathcal{F}_t (t \geq 0), P)$ will denote some underlying complete stochastic basis;
- (S, d) , an arbitrary complete separable metric space;
- $\mathcal{B}(S)$, the Borel σ -algebra on (S, d) , generated by the open balls in (S, d) ;
- $(C(S), \| \cdot \|)$, the Banach space of all real-valued bounded continuous functions on S , with the norm $\|f\| := \sup_{s \in S} |f(s)|$;
- $C_c(S)$, the subspace of $C(S)$ consisting of all those functions with compact support;
- $\mathcal{P}(S)$, the family of probability measures on $(S, \mathcal{B}(S))$;
- $C([0, \infty); S)$, the set of all continuous mappings from $[0, \infty)$ into (S, d) ;

$D([0, \infty); S)$, the set of all mappings from $[0, \infty)$ into (S, d) which are continuous from the right and possess limits from the left at every point in $(0, \infty)$;
 S^∞ , the cartesian product of countably many copies of S , with the metric $d_\infty(x, y) := \sum_{i=1}^{\infty} 2^{-i} (1 \wedge d(x_i, y_i))$ for $x = (x_i)_{i=1}^{\infty}$ and $y = (y_i)_{i=1}^{\infty}$ in S^∞ , where $a \wedge b$ denotes the minimum of a and b , two real numbers, and similarly $a \vee b$ will denote the maximum of a and b .

In the particular case of d dimensional euclidean space $\mathbb{R}^d$, $C_c^\infty(\mathbb{R}^d)$ (respectively, $C_0(\mathbb{R}^d)$; $C_0^\infty(\mathbb{R}^d)$) will denote the subspace of $C(\mathbb{R}^d)$ consisting of all those smooth (respectively, continuous; smooth) functions with compact support (respectively, with limit zero at infinity). A process, i.e. a family of random variables $X(t, \cdot) : (\Omega, \mathcal{F}, P) \rightarrow (S, \mathcal{B}(S))$, will be called continuous (respectively, cadlag) if, for each $\omega \in \Omega$, $X(\cdot, \omega)$ is continuous (respectively, right-continuous with left limits) at every time-point $t \in [0, \infty)$. We say that a process X is adapted to $\{\mathcal{F}_t : t \geq 0\}$ when $X(t, \cdot)$ is $\mathcal{F}_t$ -measurable for each $t \geq 0$.

1.1 THE SPACE $P(S)$

A sequence of $\{P_n : n \in \mathbb{Z}^+\} \subset P(S)$ will be said to converge weakly to $P \in P(S)$, and we will denote this by $P_n \Rightarrow P$, if for all

$f \in C(S)$ we have $\lim_{n \rightarrow \infty} \int f dP_n = \int f dP$. Recall the Portmanteau theorem:

Theorem: The following are equivalent:

- (i) $P_n \Rightarrow P$;
- (ii) $\lim_{n \rightarrow \infty} \int f dP_n = \int f dP$ for all bounded, uniformly continuous $f \in C(S)$;
- (iii) $\lim_{n \rightarrow \infty} P_n(F) \leq P(F)$ for all closed $F \subset S$;
- (iv) $\lim_{n \rightarrow \infty} P_n(G) \geq P(G)$ for all open $G \subset S$;
- (v) $\lim_{n \rightarrow \infty} P_n(A) = P(A)$ for all $A \in \mathcal{B}(S)$ with $P(\bar{A} \setminus A^0) = 0$,
where $\bar{A}$ denotes the closure of A in S ,
and A^0 its interior.

Proof: See Ethier and Kurtz [EK], p.108. $\square$

A set $M \subset C(S)$ will be called separating if, whenever $P, Q \in P(S)$ are such that $\int f dP = \int f dQ$ for all $f \in M$, we have $P = Q$. M will be called convergence-determining if, whenever we are given a $P \in P(S)$ and a sequence of $\{P_n : n \in \mathbb{Z}^+\} \subset P(S)$ such that $\lim_{n \rightarrow \infty} \int f dP_n = \int f dP$ holds for all $f \in M$, then $P_n \Rightarrow P$. Clearly, any convergence-determining set will be separating, since $C(S)$ itself is separating.

Proving the weak convergence of a sequence of probability measures is sometimes made easier by the existence of a small collection of nice functions which is convergence-determining. The next result provides us with one such collection. Recall that a function

$f \in C(S)$ is termed Lipschitz continuous if there is a positive number K such that for all $s_1, s_2 \in S$ we have $|f(s_1) - f(s_2)| \leq Kd(s_1, s_2)$.

1.1.2 Lemma: $C(S)$ contains a countable subset M of Lipschitz continuous functions which is convergence-determining. Further, the closure in $C(S)$ of its linear span contains $C_c(S)$. In the particular case $S = \mathbb{R}^d$, we can choose $M \subset C_c^\infty(\mathbb{R}^d)$.

Proof: Let $\{x_i\}$ denote a dense subset of S and define $f_{ij} \in C(S)$ for all $i, j \in \mathbb{Z}^+$ by $f_{ij}(s) := (2(1-jd(s, x_i)) \vee 0) \wedge 1$, i.e. f_{ij} is shaped like a truncated cone. Define

$$M := \left\{ \sum_{i,j,k} r_k f_{ij} : r_k \text{ rational}, (i,j,k) \in F \right\},$$

the maximum being taken over any finite subset of indices $F \subset (\mathbb{Z}^+)^3$.

Clearly, M is countable. Since each f_{ij} is Lipschitz continuous (with $K = 2j$) and the maximum of two Lipschitz continuous functions is again Lipschitz continuous, so are all elements of M .

We first prove that M is convergence-determining. (This part of the proof is taken from Ethier and Kurtz [EK], p.112, except for a small change.)

Let $G \subset S$ be some open set. For each $m \in \mathbb{Z}^+$, define $g_m \in M$ by fixing $r_k = 1$ for all k and selecting i and j such that both $i \vee j \leq m$ and the open ball centered at x_i of radius $1/j$, noted $B(x_i, 1/j)$, lies inside G . Then

$$\lim_{n \rightarrow \infty} P_n(G) \geq \lim_{n \rightarrow \infty} \int g_m dP_n = \int g_m dP \quad \text{for all } m.$$

the equality on the right holding true by assumption. We must show that it implies $P_n \Rightarrow P$. By (iv) of Theorem (1.1.1), it suffices to prove that $\lim_{m \rightarrow \infty} g_m(x)$ equals one if $x \in G$, zero otherwise. But for each $x \in G$ there is an integer n such that $g_m(x) = 1$ for all $m \geq n$, and $\{x \in S: g_m(x) \neq 0\} \subset G$ for every function g_m , $m \in \mathbb{Z}^+$.

In order to prove the second statement, it suffices to prove that the closure of M in $C(S)$ contains all non-negative $f \in C_c(S)$. Any such function is uniformly continuous, i.e. for all $\epsilon > 0$ there is a $\delta \in \mathbb{Z}^+$ such that for all $s_1, s_2 \in S$ we have:

$$d(s_1, s_2) < 1/\delta \text{ implies } |f(s_1) - f(s_2)| < \epsilon/2.$$

For each x_i belonging to our countable dense subset of S , we select a positive rational number r_i such that $|f(x_i) - r_i| < \epsilon/2$; thus, for all $x \in B(x_i, 1/\delta)$ we have $|f(x) - r_i| < \epsilon$.

Now, since $\bigcup_{i=1}^{\infty} B(x_i, 1/2\delta)$ covers the support of f , a compact set, there is a finite set of indices I such that $\bigcup_{i \in I} B(x_i, 1/2\delta)$ contains the support of f . Define $h := \bigvee_{i \in I} r_i f_{i\delta}$ and note that, by the last inequality above, $h(x) < f(x) + \epsilon$, for all $x \in S$. If $f(x) = 0$, this implies $|h(x) - f(x)| < \epsilon$; otherwise, there exists an index $i_0 \in I$ for which $f_{i_0\delta}(x) = 1$, so that $h(x) \geq r_{i_0} > f(x) - \epsilon$. Hence $\|h - f\| \leq \epsilon$.

For the last statement, it suffices to prove that the closure of $C_c(\mathbb{R}^d)$ in $C(\mathbb{R}^d)$ contains M . But, if f and g belong to the closure of $C_c(\mathbb{R}^d)$, so does $f \vee g = (f + g + |f - g|)/2$, since the absolute value function is approximated uniformly on bounded sets by

elements of $C_c^\infty(\mathbb{R})$ which vanish at zero. Therefore, it suffices to prove that the mappings f_{ij} belong to that closure, which may be done by elementary methods. $\square$

Let us denote by $\{f_m : m \in \mathbb{Z}^+\}$ some enumeration of the elements of M . We can now define (a version of) Varadhan's metric d_{VAR} , as follows: for any $P, Q \in P(S)$,

$$d_{\text{VAR}}(P, Q) := \sum_{m=1}^{\infty} 2^{-m} (1 \wedge |\int f_m dP - \int f_m dQ|).$$

The preceding lemma tells us that $\lim_{n \rightarrow \infty} d_{\text{VAR}}(P_n, P) = 0$ if and only if $P_n \Rightarrow P$.

(1.1.3) Theorem: $(P(S), d_{\text{VAR}})$ is a separable metric space. If S is locally compact then $(P(S), d_{\text{VAR}})$ is also complete.

Proof. The proof of separability is taken from Parthasarathy [PA].

We provide it for convenience. Let F be a dense subset of S and

put $E := \left\{ \sum_{i=1}^n r_i \delta_{x_i} : \sum_{i=1}^n r_i = 1, r_i \text{ rational}, x_i \in F \right\}$, where δ_x

denotes the Dirac measure at x . Clearly, E is dense in $P(F)$.

Given any $\varepsilon > 0$, let $\{B_i : i \in \mathbb{Z}^+\} \subset B(S)$ denote a partition of S such that $\sup\{d(x, y) : x, y \in B_i\} < \varepsilon$ for all i . Pick $s_i \in B_i \cap F$ and define $P_\varepsilon := \sum_i P(B_i) \delta_{s_i}$. For all $f \in M$, we have:

$$|\int f dP_\varepsilon - \int f dP| \leq \sup_i \left| \sup_{x \in B_i} f(x) - \inf_{x \in B_i} f(x) \right|, \text{ the left-hand side being}$$

arbitrarily small for ε small enough.

To prove the completeness, let $\{P_n : n \in \mathbb{Z}^+\} \subset P(S)$ be a d_{VAR} -Cauchy sequence and define the bounded linear functional L on the

linear span of M , noted $\text{span}(M)$, by $Lf := \lim_{n \rightarrow \infty} \int f dP_n$. There is a unique bounded linear extension of L to the closure of $\text{span}(M)$ in $C(S)$, which contains $C_c(S)$ by lemma (1.1.2) and for which the equality still holds. Since S is locally compact, by the Riesz representation theorem there is a $P \in P(S)$ such that for all $f \in C_c(S)$ we have $Lf = \int f dP$. Since $M \subset C_c(S)$ when S is locally compact, we get $\lim_{n \rightarrow \infty} d_{\text{VAR}}(P_n, P) = 0$. $\square$

We note that, for an arbitrary complete separable metric space S , the Prohorov metric d_{PRO} , defined by

$$d_{\text{PRO}}(P, Q) := \inf \{ \varepsilon > 0 : P(F) \leq Q(F^\varepsilon) + \varepsilon \text{ for all closed } F \subset S \},$$

where $F^\varepsilon = \{x \in S : \inf_{y \in F} d(x, y) < \varepsilon\}$, is such that $(P(S), d_{\text{PRO}})$ is a complete separable metric space and $\lim_{n \rightarrow \infty} d_{\text{PRO}}(P_n, P) = 0$ if and only if $P_n \Rightarrow P$ (for a proof see Ethier and Kurtz [EK], chapter 3). This implies that, for any complete separable metric space S , the space $(P(S), d_{\text{VAR}})$ is a Polish space, i.e. a separable metric space which is (only) topologically complete (see Parthasarathy [PA], chapter 2, for details).

A family of probability measures $\Phi \subset P(S)$ is called tight if for each $\varepsilon > 0$, there is a compact set $K_\varepsilon \subset S$ such that $P(K_\varepsilon) \geq 1 - \varepsilon$ for all $P \in \Phi$. We recall Prohorov's theorem, which characterizes the compact subsets of $P(S)$.

1.1.4. Theorem: $\Phi \subset P(S)$ is tight if and only if it is relatively compact.

Proof: see Ethier and Kurtz [EK], p.104.

1.2 THE SPACES $C([0, \infty): S)$ and $D([0, \infty): S)$.

Most processes considered in this work have regular paths in the sense that they can be viewed as random variables taking their values in either $C([0, \infty): S)$ or $D([0, \infty): S)$. We next review some useful facts pertaining to these spaces, beginning with the construction of the Skorohod metric $\tilde{\rho}$ on $D([0, \infty): S)$. Here again we follow Ethier and Kurtz [EK].

Let Λ be the set of continuous, strictly increasing functions from $[0, \infty)$ onto itself; then, for $\lambda \in \Lambda$ define

$$\gamma(\lambda) := 1 \wedge \left(\sup_{t \geq 0} |t - \lambda(t)| \vee \sup_{s \neq t} \left| \log \frac{\lambda(s) - \lambda(t)}{s - t} \right| \right).$$

For $\omega, \omega' \in D([0, \infty): S)$ define

$$\rho(\omega, \omega') := \int_0^\infty e^{-u} (1 \wedge \sup_{t \geq 0} d(\omega(t \wedge u), \omega'(t \wedge u))) du$$

and

$$\tilde{\rho}(\omega, \omega') := \inf_{\lambda \in \Lambda} \left(\gamma(\lambda) \vee \int_0^\infty e^{-u} (1 \wedge \sup_{t \geq 0} d(\omega(t \wedge u), \omega'(\lambda(t) \wedge u))) du \right).$$

We see that ρ is the metric of uniform convergence on all compact time-sets $[0, T]$, $T \geq 0$, whereas $\tilde{\rho}$ provides $D([0, \infty): S)$ with Stone's extension of the Skorohod topology. Intuitively, two functions ω, ω' will be close with respect to $\tilde{\rho}$ if, after making a small change in the time scale for ω' , ω and ω' are uniformly

close on all compact time-sets. For instance, two pure jump processes whose respective discontinuities are occurring "almost simultaneously" will be near each other in this sense.

One sees easily that $\tilde{\rho}$ is symmetric and transitive. If $\tilde{\rho}(\omega, \omega') = 0$, there exists, for each fixed $t > 0$, a sequence t_n converging to t such that $\lim_{n \rightarrow \infty} d(\omega(t), \omega'(t_n)) = 0$, i.e. $\omega(t) = \omega'(t)$ or $\omega'(t-)$ since ω' is right-continuous. But so is ω and thus $\omega = \omega'$. $\tilde{\rho}$ is indeed a metric.

1.2.1 Theorem: $(D([0, \infty): S), \tilde{\rho})$ is a complete separable metric space whose Borel σ -algebra is generated by the class of cylinder sets, i.e. the smallest algebra containing all the sets of the form $\tilde{X}_t^{-1}B$, for some $B \in \mathcal{B}(S)$ and some rational $t \geq 0$, where $\tilde{X}_t(\omega) := \omega(t)$.

Proof: see Ethier and Kurtz [EK], pp. 121-127.

This implies, in particular, that any S -valued cadlag process is actually a $D([0, \infty): S)$ -valued random variable. The next result illustrates the link between ρ and $\tilde{\rho}$ on $D([0, \infty): S)$.

1.2.2 Proposition: Let $\omega, \omega_1, \omega_2, \omega_3, \dots \in D([0, \infty): S)$ and denote $\omega^-(t) := \omega(t-)$; the following are equivalent:

$$(i) \quad \lim_{n \rightarrow \infty} \rho(\omega_n, \omega) = 0 \quad \text{and} \quad \omega \in C([0, \infty): S);$$

$$(ii) \quad \lim_{n \rightarrow \infty} \tilde{\rho}(\omega_n, \omega) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \rho(\omega_n, \omega_n^-) = 0.$$

Proof: (i) implies (ii) since $\tilde{\rho} \leq \rho$ and

$$\begin{aligned}
\sup_{t \geq 0} d(\omega_n(t \wedge u), \omega_n^-(t \wedge u)) &\leq \sup_{t \geq 0} d(\omega_n(t \wedge u), \omega(t \wedge u)) \\
&+ \sup_{0 < |s-t| < 1/n} d(\omega(t \wedge u), \omega(s \wedge u)) \\
&+ \sup_{s \geq 0} d(\omega(s \wedge u), \omega_n^-(s \wedge u)).
\end{aligned}$$

(ii) implies (i): fix $t > 0$ and pick any $u \in t$; then, for any $\varepsilon > 0$, there is an integer n_0 large enough that $\rho(\omega_n, \omega) < \varepsilon e^{-u}$ holds for all $n \geq n_0$. We have, for all $s \leq t$ and $\lambda \in \Lambda$,

$$\begin{aligned}
d(\omega(t \wedge u), \omega(s \wedge u)) &\leq d(\omega(t \wedge u), \omega_n(\lambda(t) \wedge u)) + \\
&d(\omega_n(\lambda(t) \wedge u), \omega_n(\lambda(s) \wedge u)) + \\
&d(\omega_n(\lambda(s) \wedge u), \omega(s \wedge u))
\end{aligned}$$

$$\text{and for some} \quad \leq 2\varepsilon + d(\omega_n(\lambda(t) \wedge u), \omega_n(\lambda(s) \wedge u)).$$

Letting s increase to t , we get

$$d(\omega(t \wedge u), \omega^-(t \wedge u)) \leq 2\varepsilon + d(\omega_n(\lambda(t) \wedge u), \omega_n^-(\lambda(t) \wedge u)).$$

Increasing n if necessary, the second part of (ii) yields that $\omega \in C([0, \infty): S)$. The first part of (i) now follows easily from this last fact. $\square$

1.2.3 Corollary: $(C([0, \infty): S), \rho)$ is a complete separable metric space.

We may now turn our attention to the compact sets in $D([0, \infty): S)$, where an analog of the Ascoli-Arzelà theorem for $C([0, \infty): S)$ holds.

For each $\omega \in D([0, \infty); S)$, each $\delta > 0$ and $T > 0$ define:

$$W'(\omega, \delta, T) := \inf_{\{t_i\}} \max_i \sup_{s, t \in [t_{i-1}, t_i)} d(\omega(s), \omega(t)), \text{ where } \{t_i\}$$

ranges over all finite collections of points of the form $0 = t_0 < t_1 < \dots < t_{n-1} < T \leq t_n$ with $n \geq 0$ and $(t_i - t_{i-1}) > \delta$ for all $i = 1, 2, \dots, n$.

1.2.4 Theorem: $\Phi \subset D([0, \infty); S)$ is relatively compact if and only if the following conditions hold:

- (i) $\tilde{X}_t \Phi$ is relatively compact for every rational $t \geq 0$;
- (ii) $\lim_{\delta \rightarrow 0} \sup_{\omega \in \Phi} W'(\omega, \delta, T) = 0$ for every $T > 0$.

Proof: see Ethier and Kurtz [EK] p.124.

In view of Prohorov's theorem (1.1.4), the preceding result provides us with a tool to prove that a sequence of probability measures on $D([0, \infty); S)$, in other words a sequence of stochastic processes with cadlag paths, has a convergent subsequence.

1.2.5 Theorem: $\{P_n : n \in \mathbb{Z}^+\} \subset P(D([0, \infty); S))$ is tight if and only if the following conditions hold:

- (i) $\{P_n \circ \tilde{X}_t^{-1} : n \in \mathbb{Z}^+\}$ is tight in $P(S)$ for every rational $t \geq 0$;

- (ii) For each $\varepsilon > 0$, $\eta > 0$, $T > 0$, there is a $\delta > 0$ and an $n_0 \in \mathbb{Z}^+$ such that

$$P_n(W'(\cdot, \delta, T) \geq \varepsilon) \leq \eta \text{ for all } n \geq n_0.$$

1.2.6 Corollary: If, in addition to the conditions (i) and (ii) of theorem (1.2.5), we have that, for each $\varepsilon > 0$, $\eta > 0$, $T > 0$, there exists an $n_0 \in \mathbb{Z}^+$ such that

$$P_n(\omega : \sup_{s \in [0, T]} d(\omega(s), \bar{\omega}(s)) \geq \varepsilon) \leq \eta \text{ for all } n \geq n_0,$$

then every limit point of $\{P_n\}$ belongs to $P(C([0, \infty); S))$.

Proof of (1.2.5) and (1.2.6): see Ethier and Kurtz [EK], pp. 129, 148. $\square$

Using the above characterization, one can derive sufficient conditions for tightness in terms of the moments of the increments of the canonical processes associated to the P_n 's. We denote the expectation of real valued functions with respect to P_n by E_n .

1.2.7 Theorem: For $\{P_n : n \in \mathbb{Z}^+\} \subset P(D([0, \infty); S))$ to be tight, it suffices that in addition to condition (i) of (1.2.5), for each $T > 0$, there exists $\beta > 0$, $C > 0$ and $\theta > 1$ such that the following conditions hold:

$$(i) \quad \lim_{h \rightarrow 0} \sup_n E_n d^\beta(\omega(h), \omega(0)) = 0;$$

(ii) For all n we have

$$E_n [d^\beta(\omega(t+h), \omega(t)) \wedge d^\beta(\omega(t), \omega(t-h))] \leq Ch^\theta$$

$$\text{for } 0 \leq t \leq T, \quad 0 \leq h \leq t.$$

Proof: see Ethier and Kurtz [EK], p. 139. $\square$

The last two results together immediately yield the celebrated criterion of Kolmogorov for tightness in $P(C([0, \infty); S))$.

1.2.8 Corollary. For $\{P_n : n \in \mathbb{Z}^+\} \subset P(C([0, \infty); S))$ to be tight, it suffices that, in addition to condition (i) of (1.2.5), for each $T > 0$ there exists $\beta > 0$, $C > 0$ and $\theta > 1$ such that, for all $n \in \mathbb{Z}^+$ and all $s, t \in [0, T]$, we have $E_n d^\beta(\omega(s), \omega(t)) \leq C|t-s|^\theta$.

Since we shall later on study sequences of diffusions with unbounded coefficients, we need a criterion for tightness which controls the size of increments of these processes prior to their random explosion times. One such criterion, due to David Aldous, may be stated as follows.

Let $X^n : (\Omega^n, F^n, F_t^n (t \geq 0), P^n) \rightarrow (S, d)$ denote a sequence of cadlag processes, each defined on its own stochastic basis and such that X^n is adapted to $\{F_t^n : t \geq 0\}$. Since S is separable, X^n may be viewed as a $D([0, \infty); S)$ -valued random variable, defining its distribution $P^n \circ (X^n)^{-1} = P^n(\{X^n\}^{-1}) \in P(D([0, \infty); S))$. We shall say that the sequence $\{X^n : n \in \mathbb{Z}^+\}$ is tight when the sequence of distributions $\{P^n \circ (X^n)^{-1} : n \in \mathbb{Z}^+\}$ is tight.

1.2.9 Theorem: For $\{X_n : n \in \mathbb{Z}^+\}$ to form a tight sequence of processes, it suffices that, in addition to condition (i) of (1.2.5), for each $\varepsilon > 0$, $\eta > 0$, $T > 0$, there exists $\delta > 0$ and $n_0 \in \mathbb{Z}^+$ such that for any sequence $(\sigma_n, \tau_n)_{n \in \mathbb{Z}^+}$ of pairs of stopping times (on $(\Omega^n, F^n, F_t^n (t \geq 0))$ respectively) with $\sigma_n \leq \tau_n \leq T$ we have

$$\sup_{n \geq n_0} P^n(d(X_{\sigma_n}^n, X_{\sigma_n}^n) \geq \eta, \tau_n < \sigma_n + \delta) \leq \varepsilon.$$

Proof: See Joffe and Métivier [JM] or Ethier and Kurtz [EK], chap.3. []

We shall not use Aldous' criterion explicitly in this thesis, but through a derived criterion of Joffe and Métivier for the tightness of sequences of semimartingales which is the object of the next section.

1.3 D-Semimartingales and the Joffe-Métivier criterion.

The material of this section is taken from the recent paper of Joffe and Métivier [JM], which the reader is referred to for the proofs of the various statements: All processes considered in this section take their values in finite-dimensional spaces, i.e. $S = \mathbb{R}^d$ equipped with the euclidean norm $|\cdot|$.

A process M is called a local martingale if there exists an increasing sequence $\{u_n: n=1,2,\dots\}$ of stopping times such that $\lim_{n \rightarrow \infty} u_n = \infty$ and for every n the stopped process $M_{t \wedge u_n}$ is an F_t -martingale. If $E|M_{t \wedge u_n}|^2 < \infty$ for all $t \geq 0$ and all n , M is said to be locally square-integrable. A process X is called a semimartingale if it can be written in the form $X = M + V$, where M is a local martingale and V is a process with paths $t \mapsto V(t, \omega)$ of finite variation on $[0, T]$ for every positive T .

The class of semimartingales possesses the double feature of containing only processes with respect to which stochastic integration may be performed (for instance, there is an extension of the Itô

formula which holds for such processes, see Métivier and Pellaumail [MP] for details) while being large enough to contain a great many Markov processes, a consequence of a Lemma of Dynkin (see Ethier and Kurtz [EK], p.162).

The class of D-semimartingales (named after Dynkin) may be described heuristically as being composed of all those semimartingales for which a formula, similar to that of Dynkin for Markov processes is known to hold. More precisely, Joffe and Métivier [JM] suggest the following definition.

1.3.1 Definition: A cadlag adapted process X , defined on a stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}_t (t \geq 0), P)$, with values in $\mathbb{R}^d$, will be called a D-semimartingale if there exists an increasing cadlag function $A(t)$, a vector space $\mathcal{C}$ of continuous real-valued functions on $\mathbb{R}^d$ and a mapping $L: (\mathcal{C} \times \mathbb{R}^d \times [0, \infty) \times \Omega) \rightarrow \mathbb{R}$ with the following properties:

(D1) The functions $x \rightarrow \phi^i(x) := x_i$ and $\phi^i \phi^j$, $i, j = 1, 2, \dots, d$ belong to $\mathcal{C}$, where $x = (x_1, x_2, \dots, x_d)$;

(D2) (i) For every $(x, t, \omega) \in \mathbb{R}^d \times [0, \infty) \times \Omega$ the mapping $\phi \rightarrow L(\phi, x, t, \omega)$ is a linear functional on $\mathcal{C}$ and $L(\phi, \cdot, t, \omega) \in \mathcal{C}$;

(ii) For every $\phi \in \mathcal{C}$, $(x, t, \omega) \rightarrow L(\phi, x, t, \omega)$ is $\mathcal{B}(\mathbb{R}^d) \otimes \underline{\mathcal{P}}$ -measurable, where $\underline{\mathcal{P}}$ is the predictable σ -algebra on $[0, \infty) \times \Omega$, i.e. $\underline{\mathcal{P}}$ is generated by the sets of the form $(s, t] \times F$, where $F \in \mathcal{F}_s$ and s, t are arbitrary;

(D3) For every $\phi \in C$, the process M^ϕ defined by $M^\phi(t; \omega) := \phi(X_t(\omega)) - \phi(X_0(\omega)) - \int_0^t L(\phi, X_{s-}(\omega), S, \omega) dA_s$ is a locally square-integrable martingale on $(\Omega, F, F_t (t \geq 0), P)$.

This last condition is precisely a formula like Dynkin's.

Define the vector $\underline{\beta} = (\beta^i)_{i=1,2,\dots,d}$ and the matrix $\underline{\alpha} = (\alpha^{ij})_{i,j=1,2,\dots,d}$ by

$$\beta^i(x, t, \omega) := L(\phi^i, x, t, \omega)$$

and $\alpha^{ij}(x, t, \omega) := L(\phi^i \phi^j, x, t, \omega) - x_i \beta^j(x, t, \omega) - x_j \beta^i(x, t, \omega),$

respectively called the local coefficients of first and second order.

They correspond to the drift and diffusion coefficients of the generator in the case where X is actually a Markov process. We can now state the extension of Rebolledo's criterion for tightness of a sequence of martingales which is due to Joffe and Métivier.

1.3.2 Theorem: Let $X^m: (\Omega^m, F^m, F_t^m (t \geq 0), P^m) \subset \mathbb{R}^d$, $m=1,2,3,\dots$, be a sequence of D-semimartingales, with common C and associated operators L^m , functions A^m , vectors $\underline{\beta}_m$ and matrices $\underline{\alpha}_m$ as defined in (1.3.1). Then the sequence $\{X^m: m \in \mathbb{Z}^+\}$ is tight provided the following conditions hold:

$$(i) \sup_m E |X^m(0)|^2 < \infty;$$

(ii) there is a $K > 0$ and a sequence of positive adapted processes $\{C_t^m: t \geq 0\}$ such that for every $m \in \mathbb{Z}^+$, $x \in \mathbb{R}^d$, $\omega_m \in \Omega^m$:

$$a) \quad |\beta_m(x, t, \omega_m)|^2 + \text{trace } \alpha_m(x, t, \omega_m) \leq K(C_t^m(\omega_m) + |x|^2);$$

b) for every $T > 0$,

$$\sup_m \sup_{t \in [0, T]} EC_t^m < \infty \text{ and } \lim_{\rho \rightarrow \infty} \sup_m P^m(\sup_{t \in [0, T]} C_t^m \geq \rho) = 0;$$

iii) There exists a positive function γ on $[0, \infty)$ and a decreasing sequence of numbers $\{\rho_m : m=1, 2, \dots\}$ such that $\lim_{t \rightarrow 0} \gamma(t) = 0$, $\lim_{m \rightarrow \infty} \rho_m = 0$ and for all $0 < s < t$ and all m :

$$(A^m(t) - A^m(s)) \leq \gamma(t-s) + \rho_m.$$

Further, if we set $M_t^m := X_t^m - X_0^m - \int_0^t \beta_m(X_{s-}^m, s, \cdot) dA_s^m(\cdot)$, then for each $T > 0$ there exists a constant K_T and m_0 such that, for all $m \geq m_0$:

$$iv) \quad E(\sup_{t \in [0, T]} |X_t^m|^2) \leq K_T(1 + E|X_0^m|^2);$$

and

$$v) \quad E(\sup_{t \in [0, T]} |M_t^m|^2) \leq K_T(1 + E|X_0^m|^2).$$

Proof: see Joffe and Métivier [JM]. $\square$

When applying this powerful result in this thesis, the functions A^m will be equal to the identity, i.e. $A^m(t) = t$, so condition (iii) above will not interest us. Condition (i) above clearly implies condition (i) of theorem (1.2.5). The important condition here of course is (ii), which can be viewed as an extended moment condition of the type in theorem (1.2.7). As a bonus, Joffe and Métivier get the inequalities (iv) and (v), which are actually derived first and used in verifying that the condition of Aldous in theorem (1.2.9) above holds.

CHAPTER 2

EXCHANGEABLE SYSTEMS OF PARTICLES

After introducing the notion of exchangeability for sequences of processes, we give some weak conditions for a finite Markovian system of processes to exhibit exchangeability: first, in the general case (section 1) in terms of the transition kernel; then, in section 3, in the case of multidimensional diffusions, in terms of the coefficients of the generator of the process. In section 4, we derive a law of large numbers for the empirical measure of the limiting infinite system, proving at the same time the existence of solutions to a large class of martingale problems for measure-valued diffusions.

2.1 Exchangeable sequences of processes

In this section, we define exchangeability for sequences of stochastic processes and characterize them in terms of a weaker form of exchangeability in the case of Markov processes. Here again, let $(\Omega, \mathcal{F}, \mathcal{F}_t (t \geq 0), P)$ denote some underlying complete stochastic basis and (S, d) a complete separable metric space; then

$(\tilde{\Omega}, \tilde{\mathcal{F}})$ will denote some measurable space;

$(\tilde{\Omega}^n, \tilde{\mathcal{F}}^n), (\tilde{\Omega}^\infty, \tilde{\mathcal{F}}^\infty)$, the n -fold and countably infinite product spaces with their respective product σ -algebras generated by products of elements of $\tilde{\mathcal{F}}$;

$\text{Perm}(n)$, the group of permutations on $\{1, 2, \dots, n\}$, for each $n \in \mathbb{Z}^+$;

- $\text{Perm}(\infty)$, the group of finite permutations on $\mathbb{Z}^+$, i.e. those mappings $\pi : \mathbb{Z}^+ \rightarrow \mathbb{Z}^+$ such that $\pi(n) = n$ for all but finitely many n (occasionally, we write πn for $\pi(n)$);
- $Q(n)$, the set of mappings $\Pi : \tilde{\Omega}^n \rightarrow \tilde{\Omega}^n$ satisfying $\Pi(\omega_1, \omega_2, \dots, \omega_n) := (\omega_{\pi_1}, \omega_{\pi_2}, \dots, \omega_{\pi_n})$ for some $\pi \in \text{Perm}(n)$;
- $Q(\infty)$, the set of mappings $\Pi : \tilde{\Omega}^\infty \rightarrow \tilde{\Omega}^\infty$ satisfying $\Pi(\omega_1, \omega_2, \omega_3, \dots) := (\omega_{\pi_1}, \omega_{\pi_2}, \omega_{\pi_3}, \dots)$ for some $\pi \in \text{Perm}(\infty)$;
- $\tilde{Q}(\tilde{n})$, the set of mappings $\tilde{\Pi} : D([0, \infty) : S^n) \rightarrow D([0, \infty) : S^n)$ defined by $\tilde{\Pi}(\omega)(t) := \Pi(\omega(t))$, for some $\Pi \in Q(n)$ and all $t \geq 0$ (of course, here $\tilde{\Omega} = S$ and $\tilde{F} = \mathcal{B}(S)$);
- similarly with ∞ replacing n everywhere;
- $\tilde{X}_{(t_1, t_2, \dots, t_k)}$, the mapping from $D([0, \infty) : S) \rightarrow S^k$, for each $k \in \mathbb{Z}^+$ and $(t_1, t_2, \dots, t_k) \in [0, \infty)^k$, defined by $\tilde{X}_{(t_1, t_2, \dots, t_k)}(\omega) := (\omega(t_1), \omega(t_2), \dots, \omega(t_k))$;
- $M(S)$, the set of all Borel-measurable functions from S to $\mathbb{R}$;
- $\overline{B}(S) \subset M(S)$, the bounded functions from S to $\mathbb{R}$;
- $\mathcal{B}_t$, the σ -algebra on $D([0, \infty) : S)$ generated by all the sets of the form $\tilde{X}_t^{-1}B$, where $B \in \mathcal{B}(S)$ and $t \in [0, \infty)$.

An n -exchangeable sequence is a random variable $Z = (Z_1, Z_2, \dots, Z_n) : (\Omega, \mathcal{F}, P) \rightarrow (\tilde{\Omega}^n, \tilde{F}^n)$ which satisfies $P \circ Z^{-1}(B) = P \circ Z^{-1}(\Pi B)$ for all $B \in \tilde{F}^n$ and all $\Pi \in Q(n)$. When n is left unspecified we sometimes call Z a finite exchangeable sequence. Similarly, an

infinite exchangeable sequence is a random variable $Z = (Z_1, Z_2, Z_3, \dots)$: $(\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\tilde{\Omega}^\infty, \tilde{\mathcal{F}}^\infty)$ which satisfies $\mathbb{P} \circ Z^{-1}(B) = \mathbb{P} \circ Z^{-1}(\Pi B)$ for all $B \in \tilde{\mathcal{F}}^\infty$ and all $\Pi \in Q(\infty)$. A classical construction of such sequences is provided by sampling schemes.

2.1.1 Example. Let $\{Y_n : n=1, 2, \dots, m\}$ denote a collection of $(\tilde{\Omega}, \tilde{\mathcal{F}})$ -valued random variables and $\{u(i) : i \in \underline{Z}^+\}$ a sequence of random variables, independent of each other and of the X_n 's and each distributed uniformly over $\{1, 2, \dots, m\}$. All the processes are built on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then, $Z_n := Y_{u(n)}$ defines an infinite exchangeable sequence on $(\Omega, \mathcal{F}, \mathbb{P})$. Replacing $\underline{Z}^+$ by $\{1, 2, \dots, k\}$ above yields the k -exchangeable case. To prove this, simply notice that

$$P(Z_n \in B_n, n=1, 2, \dots, k) = \frac{1}{m^k} \sum P(Y_{u_n} \in B_n, n=1, 2, \dots, k | u(n)=u_n, n=1, 2, \dots, k)$$

where the sum runs over all vectors $(u_1, u_2, \dots, u_k) \in \{1, 2, \dots, m\}^k$.

For more on exchangeability of sequences of random variables, the readers should look in the survey of Aldous [AL]. Our interest in the notion of exchangeability stems from the following result.

2.1.2 Theorem: Assume S is locally compact. If $Z : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (S^\infty, \mathcal{B}(S^\infty))$ is an infinite exchangeable sequence, there exists a $P(S)$ -valued random variable μ such that $P(\mu_m := \frac{1}{m} \sum_{n=1}^m \delta_{Z_n} \Rightarrow \mu) = 1$.

Proof: $\{\mu_m(\omega) : m \in \underline{Z}^+\}$ forms a Cauchy sequence in $(P(S), dv)$ if and only if $\{\int f d\mu_m(\omega) : m \in \underline{Z}^+\}$ form a real Cauchy sequence, for every $f \in M$, the countable set of bounded functions built in lemma (1.1.2).

But this is the case for almost all $\omega \in \Omega$ by De Finetti's strong law of large numbers (see Billingsley [BI] or Aldous [AL]) applied to the bounded exchangeable sequence $\{f(Z_n) : n \in \mathbb{Z}^+\}$. Thus, for each ω outside a set of measure zero, there exists a $\mu(\omega) \in P(S)$ such that $\mu_m(\omega) \Rightarrow \mu(\omega)$, by theorem (1.1.3). The measurability of μ follows from the completeness of the σ -algebra $\mathcal{F}$ and the fact that pointwise limits of measurable functions taking values in an arbitrary metric space are again measurable. $\square$

For instance, one sees at once that, in the case of example (2.1.1), we can characterize the limit, namely $\mu = \frac{1}{m} \sum_{n=1}^m \delta_{Y_n}$.

We are now in position to present some results on exchangeable sequences of processes.

A sequence $\{Z_n : n \in \mathbb{Z}^+\}$ of $D([0, \infty); S)$ -valued random variables, built on a common probability space $(\Omega, \mathcal{F}, P)$, will be called k-fold exchangeable, if, for every k-tuple $(t_1, t_2, \dots, t_k)$ such that $0 \leq t_1 < t_2 < \dots < t_k < \infty$, the sequence $\{\tilde{X}_{(t_1, t_2, \dots, t_k)}(Z_n) : n \in \mathbb{Z}^+\}$ of S^k -valued random variables is exchangeable. A similar definition for finite sequences is obtained by substituting a finite set to $\mathbb{Z}^+$ in the above definition. In the case where S is a vector space, such a sequence will be said to have exchangeable increments if, for every pair $(s, t) \in [0, \infty)^2$, $\{Z_n(t) - Z_n(s) : n \in \mathbb{Z}^+\}$ of S -valued random variables is exchangeable. Note that our sequences of processes with exchangeable increments are not the same as those mentioned in Aldous [AL]. Here, exchangeability is in n , not in t , the interest being

in modelling systems of indistinguishable particles and not in the particular sample path behavior of an individual particle. We have the following

2.1.3 Proposition: For a sequence $\{Z_n : n \in \mathbb{Z}^+\}$ of $D([0, \infty): S)$ -valued random variables built on a common probability space, "exchangeability" is equivalent to "k-fold exchangeability for all $k \in \mathbb{Z}^+$ "; "2-fold exchangeability" implies "exchangeability of increments".

Proof: By theorem (1.2.1), $B(D([0, \infty): S))$ is generated by the sets $\tilde{X}_{(t_1, t_2, \dots, t_k)}^{-1}(B)$ where $B \in B(S^k)$, $k \in \mathbb{Z}^+$. Since S is separable, so is $D([0, \infty): S)$ and that implies $B(D([0, \infty): S)^i) = B(D([0, \infty): S))^i$ for each $i \in \mathbb{Z}^+$, where of course the first product indicates the cartesian product while the second one denotes the smallest σ -algebra containing all i -fold products of Borel sets.

In order to prove that, for all choices of $i \in \mathbb{Z}^+$, $\tilde{\Pi}_i \in Q(i)$, $B \in B(D([0, \infty): S)^i)$, we have $P \circ (Z_1, Z_2, \dots, Z_i)^{-1}(\tilde{\Pi}_i B) = P \circ (Z_1, Z_2, \dots, Z_i)^{-1}(B)$, it therefore suffices to consider sets B of the form $B = \bigcap_{j=1}^i \tilde{X}_{(t_{j1}, t_{j2}, \dots, t_{jk})}^{-1}(B_j)$ for each value of $k \in \mathbb{Z}^+$ and $B_j \in B(S^k)$, $j=1, 2, \dots, i$. So we must show that

$$P(\tilde{X}_{(t_{j1}, t_{j2}, \dots, t_{jk})}^{-1}(Z_{\pi_j}) \in B_j, j=1, 2, \dots, i) = P(\tilde{X}_{(t_{j1}, t_{j2}, \dots, t_{jk})}^{-1}(Z_j) \in B_j, j=1, 2, \dots, i)$$

for all choices of $i \in \mathbb{Z}^+$, $\pi \in \text{Perm}(i)$, $k \in \mathbb{Z}^+$, $B_j \in B(S^k)$ and $0 \leq t_{j1} < t_{j2} < \dots < t_{jk} < \infty$. But this is exactly k -fold exchangeability for every value of k .

The fact that 2-fold exchangeability implies exchangeability of the increments follows in a similar fashion from the continuity of the mapping $(x,y) \in S^2 \rightarrow (x-y)$.

The same result holds for finite sequences, with the appropriate change. We next seek a partial converse to this proposition, in the Markovian case. To this end, we need the following

2.1.4 Definition: A mapping $q : [0, \infty) \times \tilde{\Omega} \times [0, \infty) \times \tilde{F} \rightarrow [0, 1]$ is called a transition function if it satisfies the following three conditions:

- (i) $q(s, x; t, \cdot)$ is a probability measure on $(\tilde{\Omega}, \tilde{F})$ for every $0 \leq s < t < \infty$ and $x \in \tilde{\Omega}$;
- (ii) $q(s, \cdot; t, B)$ is $\tilde{F}$ -measurable for every $0 \leq s < t < \infty$ and every $B \in \tilde{F}$;
- (iii) $q(s, x; t, B) = \int q(u, y; t, B) q(s, x; u, dy)$, the Chapman-Kolmogorov equation.

We shall say that an F_t -adapted process $Y : (\Omega, F, F_t (t \geq 0), P) \rightarrow (\tilde{\Omega}, \tilde{F})$ is a Markov process if, for every $0 \leq s < t < \infty$ and every set $F \in \tilde{F}$, we have $P(Y_t \in F | F_s) = P(Y_t \in F | Y_s)$. A Markov process Y with transition function q will then just be any such process satisfying $P(Y_t \in F | F_s) = q(s, Y_s; t, F)$, for all $0 \leq s < t < \infty$ and all $F \in \tilde{F}$.

2.1.5 Proposition: If $Z = (Z_1, Z_2, \dots, Z_n)$ is a sequence of $D([0, \infty): S)$ -valued random variables, built on the same stochastic basis

$(\Omega, \mathcal{F}, \mathcal{F}_t (t \geq 0), P)$ and such that Z is a Markov process with transition function q on $[0, \infty) \times \Omega \times [0, \infty) \times S^n$, then Z is exchangeable provided it is 1-fold exchangeable and $q(s, y; t, A) = q(s, \Pi y; t, \Pi A)$ for every choice of $\Pi \in Q(n)$, $A \in \mathcal{B}(S^n)$, $y \in S^n$ and $0 \leq s < t < \infty$.

$$\text{Proof: } P(\tilde{X}_{(t_1, t_2, \dots, t_k)}(Z_m) \in \bigcap_{i=1}^k A_i^m, m=1, 2, \dots, n)$$

$$= P(Z(t_1) \in \bigcap_{m=1}^n A_i^m, i=1, 2, \dots, k)$$

$$= \int \dots \int P(Z(t_1) \in dy_1) \prod_{i=1}^{k-1} q(t_1, y_i; t_{i+1}, dy_{i+1}),$$

$$\bigcap_{i=1}^k \bigcap_{m=1}^n A_i^m$$

the equalities still holding when A_i^m is replaced everywhere by $A_i^{\pi m}$, for some $\pi \in \text{Perm}(n)$. But these two sets of equalities are the same since by assumption $P(Z(t_1) \in d\Pi y_1) = P(Z(t_1) \in dy_1)$ and $q(t_i, \Pi y_i; t_{i+1}, d\Pi y_{i+1}) = q(t_i, y_i; t_{i+1}, dy_{i+1})$ hold for all $\Pi \in Q(n)$. Hence, Z is k -exchangeable for every $k \in \mathbb{Z}^+$, which implies that Z is exchangeable by proposition (2.1.3).

The corresponding result for infinite sequences of processes is somewhat less interesting.

2.1.6 Corollary: If $Z = (Z_1, Z_2, Z_3, \dots)$ is a sequence of $D([0, \infty); S)$ -valued random variables, built on the same stochastic basis and such that $(Z_1, Z_2, \dots, Z_m)$ is a Markov process with transition function q_m on $[0, \infty) \times \Omega \times [0, \infty) \times S^m$ for every $m \in \mathbb{Z}^+$, then Z is exchangeable provided it is 1-fold exchangeable and

$q_m(s, y; t, A_m) = q_m(s, \prod_{m=1}^m y_m; t, \prod_{m=1}^m A_m)$ for every choice of $m \in \mathbb{Z}^+$,

$\pi_m \in Q(m)$, $A_m \in B(S^m)$, $y_m \in S^m$ and $0 \leq s < t < \infty$.

What can be said about an analog of theorem (2.1.2) when the locally compact space S is replaced by $D([0, \infty): S)$? In general, it is asking for too much to expect weak convergence with probability one. However, for the finite-dimensional distributions, we can assert the following:

2.1.7. Proposition: Assume S is locally compact. For any k -fold exchangeable sequence of $D([0, \infty): S)$ -valued random variables $\{Z_n: n \in \mathbb{Z}^+\}$, there exists, for each choice of $0 \leq t_1 < t_2 < \dots < t_k < \infty$, a $P(S^k)$ -valued random variable $\mu(t_1, t_2, \dots, t_k)$ such that

$$P\left(\frac{1}{m} \sum_{n=1}^m \delta_{X_{t_1, t_2, \dots, t_k}(Z_n)} = \mu(t_1, t_2, \dots, t_k)\right) = 1.$$

Proof: Since S is locally compact, so is S^k and hence theorem (2.1.2) applies.

As a consequence, when S is locally compact and the sequence is 1-fold exchangeable, all finite-dimensional distributions of the $D([0, \infty): P(S))$ -valued random variables $\frac{1}{m} \sum_{n=1}^m \delta_{Z_n}$ converge to unique limits (which unfortunately may well not be the finite-dimensional distributions of any $D([0, \infty): P(S))$ -valued random variable).

In the next sections, we turn our attention to a smaller class of exchangeable processes for which more can be said about the asymptotic behavior of the empirical measure $\frac{1}{m} \sum_{n=1}^m \delta_{Z_n}$.

2.2. Martingale Problems: Uniqueness and Duality

One of the most powerful methods for proving weak convergence of sequences of probability measures on $D([0, \infty); S)$ and characterizing the limit is the martingale problem formulation, due to Stroock and Varadhan [SV]. We summarize here some of the general theory associated with such problems, as expounded in the book of Ethier and Kurtz [EK], specializing some of their statements when their full generality is not needed.

We recall that, given a subset $A \subset \bar{B}(S) \times \bar{B}(S)$, where $\bar{B}(S)$ denotes the space of bounded measurable functions from S to $\mathbb{R}$, and a starting measure $\mu \in P(S)$, we say that $P \in P(D([0, \infty); S))$ is a solution to the martingale problem for (A, μ) if the following hold:

- (i) $P \circ \tilde{X}_{(0)}^{-1} = \mu$;
- (ii) $f \circ X_{(t)} - \int_0^t g \circ \tilde{X}_{(s)} ds$ is a $\mathcal{B}_t$ -martingale with respect to P for each $(f, g) \in A$.

The martingale problem for (A, μ) is termed well-posed when it possesses one and only one solution. The problem of existence of solutions for those martingale problems of interest to us will be taken up in a later section of this chapter and again in chapter 3. For now, we recall some results concerning the uniqueness of solutions, following the treatment given in Ethier and Kurtz [EK].

2.2.1 Theorem: Suppose that, for each $\mu \in P(S)$, any two solutions P_1, P_2 of the martingale problem for (A, μ) satisfy the equation

$P_1 \circ \tilde{X}_t^{-1} = P_2 \circ \tilde{X}_t^{-1}$ for every $t \geq 0$; then $P_1 = P_2$ and $\tilde{X}$ is a Markov process with respect to $\mathcal{B}_t$ ($t \geq 0$), given any starting point μ . Further, if $A \in \mathcal{C}(S) \times \bar{\mathcal{B}}(S)$ and for each $x \in S$ there is a solution P_{δ_x} to the martingale problem for (A, δ_x) such that the mapping $x \rightarrow P_{\delta_x}$ is Borel-measurable, then $\tilde{X}$ is a strong Markov process with respect to $\mathcal{B}_t$ ($t \geq 0$).

Proof: see Ethier and Kurtz [EK], pp. 184 to 187.

In order to prove the uniqueness of solution to the martingale problem for (A, μ) , it thus suffices to prove the uniqueness of the one-dimensional distributions. A new and powerful tool for proving this is the method of duality originating in the work of Vasershtein and Spitzer, which we now briefly explain. Heuristically, it consists in the construction of another process, valued in another space, but possessing the feature of being a solution to a martingale problem for an operator which is formally adjoint to A , in a sense made precise below.

2.2.2 Definition: Let S_1, S_2 denote separable metric spaces, S_1 complete; let $A_i \in \bar{\mathcal{B}}(S_i) \times \bar{\mathcal{B}}(S_i)$, $i=1,2$; $F \in M(S_1 \times S_2)$; $\mu_i \in \mathcal{P}(S_i)$, $i=1,2$. The martingale problems for (A_1, μ_1) and (A_2, μ_2) are called dual with respect to F if, for each solution P_1 for (A_1, μ_1) and P_2 for (A_2, μ_2) , we have, for all $t \geq 0$:

$$(i) \int E_1 |F(\tilde{X}_{(t)}, y)| d\mu_2(y) < \infty;$$

$$(ii) \int E_2 |F(x, \tilde{Y}_{(t)})| d\mu_1(x) < \infty;$$

$$(iii) \int E_1 F(\tilde{X}_{(t)}, y) d\mu_2(y) = \int E_2 F(x, \tilde{Y}_{(t)}) d\mu_1(x).$$

Note: When $\tilde{X}$ and $\tilde{Y}$ are built on a common probability space and independent, (iii) becomes

$$(iii)' \quad EF(\tilde{X}_{(t)}, \tilde{Y}_{(0)}) = EF(\tilde{X}_{(0)}, \tilde{Y}_{(t)}).$$

2.2.3 Theorem: If (A_1, μ) and (A_2, δ_y) are dual with respect to F for every $\mu \in P(S_1)$ with compact support and every $y \in S_2$; and if $\{F(\cdot, y) : y \in S_2\}$ is separating on $P(S_1)$; then the existence of a solution to the martingale problems for (A_2, δ_y) , for every $y \in S_2$, implies the uniqueness of solution to the martingale problem for (A_1, μ) , for every $\mu \in P(S_1)$.

Proof: see Ethier and Kurtz [EK], p. 190.

Finding a dual process is, in many cases where one is dealing with Markov processes, a matter of studying the form of the generator. The following theorem, a special case of a general result of Dawson and Kurtz [DK], makes this point clearly.

2.2.4 Theorem: Suppose that X and Y are independent cadlag processes such that, for $F, G \in M(S_1 \times S_2)$ we have

- (i) $F(X_s, Y_t) - \int_0^s G(X_r, Y_t) dr$ is an F_s^X -martingale, for each $t \geq 0$;
- (ii) $F(X_s, Y_t) - \int_0^t G(X_s, Y_r) dr$ is an F_t^Y -martingale, for each $s \geq 0$;
- (iii) $\sup_{s, t \in [0, \infty)} E|F(X_s, Y_t)| < \infty$ and $\sup_{s, t \in [0, \infty)} E|G(X_s, Y_t)| < \infty$.

Then

(D) $EF(X_t, Y_0) = EF(X_0, Y_t)$ for every $t \geq 0$.

Proof: see Dawson and Kurtz [DK] or Ethier and Kurtz [EK], pp. 192 to 195.

2.3 Exchangeable sequences of diffusion

Our aim is now to characterize those finite systems of particles which exhibit exchangeability of trajectories, under the assumption that the motion is continuous and Markovian, in terms of the coefficients of the generator appearing in the martingale formulation. Then we characterize those exchangeable systems, all of whose subsystems (formed by observing only a subset of the particles) possess the Markov property. The motion of such a collection of particles will have the following features: the particles are indistinguishable, each subcollection moves continuously through space, starting afresh at random times regardless of the position of the rest of the particles at that time. This does not imply that the motions of the particles are independent of one another for all times. The evolution will often be coherent, in a sense made precise below (see pp. 37-38).

To this end, let $S = \mathbb{R}^d$ and m denote the number of particles in some finite system. Put $n = dm$. An $n \times n$ matrix $\underline{a}$ will sometimes be written in the form of a block matrix $\underline{a} = (a_{ij})$, $i, j = 1, 2, \dots, m$, where a_{ij} is a $d \times d$ matrix, in the obvious way. Similarly, a vector $\underline{b} \in \mathbb{R}^n$, written as a column, will satisfy $\underline{b}^T = (b_i^T)$, $i = 1, 2, \dots, m$, for $b_i \in \mathbb{R}^d$. We shall write likewise

the gradient operator ∇ acting on $C_c^\infty(\mathbb{R}^n)$ as $\nabla^T = (\nabla_i^T)$, $i = 1, 2, \dots, m$, where ∇_i is the gradient on $C_c^\infty(\mathbb{R}^d)$, acting on the i^{th} block of d coordinates of $x \in \mathbb{R}^n$, i.e. on variables $x_{di+1}, x_{di+2}, \dots, x_{di+d}$. Finally, $|\cdot|$ will denote the euclidean norm, sometimes applied to matrices in the usual way.

The problem of existence and uniqueness of a solution to the martingale problem for a large class of differential operators has been extensively researched (see Stroock and Varadhan [SV] or Ethier and Kurtz [EK] for more detail).

2.3.1 Theorem: Given a continuous function $\underline{b}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ and a continuous function $\underline{a}$ from $\mathbb{R}^n$ to the space of real $n \times n$ symmetric positive definite matrices, such that, for some $k > 0$, $|\underline{a}(x)| \leq k(1 + |x|^2)$ and $x \cdot \underline{b}(x) \leq k(1 + |x|^2)$ for each $x \in \mathbb{R}^n$; then, the martingale problem for $A = \{(f, Lf) : f \in C_c^\infty(\mathbb{R}^n)\}$, where

$$L := \frac{1}{2}(\underline{a}\nabla)^T \nabla + \underline{b}^T \nabla,$$

is well-posed and its unique solution P belongs to $P(C([0, \infty); S^m))$.

Proof: see Stroock and Varadhan [SV], p. 255.

Part of the conclusion of this last theorem is that there is no explosion in finite time, i.e. the process $\tilde{X}_t$ belongs to $\mathbb{R}^n$ for all $t \geq 0$ and all paths. Actually, more is true under the stated conditions, namely:

2.3.2 Corollary: Under the conditions of theorem (2.3.1), there exists a constant $C > 0$ such that, for all $t \geq 0$, we have

$$E|X_t|^2 \leq e^{Ct} (1 + E|\tilde{X}_0|^2).$$

Proof: Letting $\tau_h := \inf \{t > 0 : |\tilde{X}_t| \geq h\}$ and $\phi \in C_c^\infty(\mathbb{R}^n)$ such that $\phi(x) = |x|^2$ whenever $|x| \leq h$, then

$$\begin{aligned} E\phi(\tilde{X}_{t \wedge \tau_h}) - E\phi(\tilde{X}_0) &= E \int_0^{t \wedge \tau_h} (\text{trace}(\underline{a}(\tilde{X}_s)) + 2\tilde{X}_s^T \underline{b}(\tilde{X}_s)) ds \\ &\leq (n+2)K \int_0^{t \wedge \tau_h} (1 + E|\tilde{X}_s|^2) ds. \end{aligned}$$

So

$$\begin{aligned} 1 + E|\tilde{X}_{t \wedge \tau_h}|^2 &\leq 1 + E|\tilde{X}_0|^2 + (n+2)K \int_0^t (1 + E|\tilde{X}_{s \wedge \tau_h}|^2) ds \\ &\leq (1 + E|\tilde{X}_0|^2) e^{(n+2)Kt}. \end{aligned}$$

Letting $h \rightarrow \infty$ yields $\tau_h \rightarrow \infty$ and the stated result for

$$C = (n+2)K.$$

Our first result describes the relations satisfied by the coefficients of the generator of a finite exchangeable sequence of diffusions in the finite-dimensional case. The family $\{P_y : y \in S^m\}$ of solutions to the martingale problems for L , started at $y \in S^m$, mentioned in theorem (2.3.1) above, will be called exchangeable if it satisfies $P_{\Pi y} = P_y \circ \Pi^{-1}$, for every $y \in S^m$, every $\Pi \in \tilde{\phi}(m)$ and corresponding $\Pi \in \phi(m)$.

2.3.3 Theorem: Let $\underline{a}$ and $\underline{b}$ satisfy the conditions of theorem (2.3.1). In addition, assume that $|\underline{b}(x)| \leq K(1+|x|^2)$ for each $x \in \mathbb{R}^n$. Then, the solutions $\{P_y : y \in S^m\}$ to the martingale problems for operator $L := \frac{1}{2}(\underline{a}\nabla)^T \nabla + \underline{b}^T \nabla = \frac{1}{2} \sum_{i,j=1}^m (a_{ij} \nabla_i)^T \nabla_j + \sum_{i=1}^m b_i^T \nabla_i$

forms an exchangeable family (i.e. a system of m -exchangeable d -dimensional processes) if and only if the following conditions hold:

$$2.3.4 \quad \begin{cases} b_i(\Pi y) = b_{\pi_i}(y), & \text{for all } i=1,2,\dots,m, y \in S^m, \Pi \in \phi(m) \text{ and} \\ & \text{corresponding } \pi \in \text{Perm}(m); \\ a_{ij}(\Pi y) = a_{\pi_i \pi_j}(y), & \text{for all } i,j=1,2,\dots,m, y \in S^m, \Pi \in \phi(m) \\ & \text{and corresponding } \pi \in \text{Perm}(m). \end{cases}$$

2.3.5 Remark. If we put $\mu_m := \frac{1}{m} \sum_{i=1}^m \delta_{y_i} \in \mathcal{P}(S)$ and define

$$2.3.6 \quad \begin{cases} b(y_1, \mu_m) := b_1(y); \\ \sigma(y_1, \mu_m) := a_{11}(y); \\ \rho(y_1, y_2, \mu_m) := a_{12}(y); \end{cases}$$

then any pair $(\underline{a}, \underline{b})$ satisfying (2.3.4) corresponds in a unique way to some triple (b, σ, ρ) through equations (2.3.6) in such a way that the following holds everywhere:

$$2.3.7 \quad \begin{cases} b_i(y) = b(y_i, \mu_m), & i = 1, 2, \dots, m; \\ a_{ii}(y) = \sigma(y_i, \mu_m), & i = 1, 2, \dots, m; \\ a_{ij}(y) = \rho(y_i, y_j, \mu_m), & i, j = 1, 2, \dots, m, i \neq j. \end{cases}$$

Proof of theorem (2.3.3): To avoid a mess of subindices, we shall write, f for $(f_k)_{k=1}^m : \mathbb{R}^n \rightarrow \mathbb{R}^d$, where $f_k \in C_c^\infty(\mathbb{R}^n)$, or f for $(f_{kl})_{k,l=1}^m : \mathbb{R}^n \rightarrow$ the space of $d \times d$ matrices, respectively Lf for $(Lf_k)_{k=1}^m$ or Lf for $(Lf_{kl})_{k,l=1}^m$. Similarly, expectations should be taken componentwise when necessary.

- (Necessity). For any $f_1, f_2 \in C_c^\infty(S^m)$ and any fixed $t \in [0, \infty)$,

$$E_{\Pi_Y} f_1(\tilde{X}(t)) - E_{\Pi_Y} f_1(\tilde{X}(0)) = E_{\Pi_Y} \int_0^t L f_1(\tilde{X}(s)) ds, \text{ for all } \pi \in \phi(m),$$

and

$$E_Y f_2(\tilde{X}(t)) - E_Y f_2(\tilde{X}(0)) = E_Y \int_0^t L f_2(\tilde{X}(s)) ds.$$

(i) Put $f_1(x) = x_i$, $f_2(x) = x_{\pi i}$ in the above equations to

get:

$$E_{\Pi_Y} \tilde{X}_i(t) - y_{\pi i} = E_{\Pi_Y} \int_0^t b_i(\tilde{X}(s)) ds,$$

$$E_Y \tilde{X}_{\pi i}(t) - y_{\pi i} = E_Y \int_0^t b_{\pi i}(\tilde{X}(s)) ds,$$

after a stopping argument of the type used in corollary (2.3.2), and using our new condition on $\underline{b}$. The (1-fold) exchangeability implies that the two left-hand sides are equal. Therefore, using Fubini's theorem we get

$$\int_0^t (E_{\Pi_Y} b_i(\tilde{X}(s)) - E_Y b_{\pi i}(\tilde{X}(s))) ds = 0 \text{ for all } t \geq 0.$$

Hence $E_{\Pi_Y} b_i(\tilde{X}(s)) = E_Y b_{\pi i}(\tilde{X}(s))$ for almost every $s \in [0, \infty)$.

But, since $\underline{b}$ and $\tilde{X}$ are continuous, the bounded convergence theorem (applicable here again because of the extra condition on the growth of $\underline{b}$ and corollary (2.3.2)) implies

$$b_i(\Pi y) = b_{\pi i}(y).$$

(ii) Now put $f_1(x) = x_i x_j^T$, $f_2(x) = x_{\pi i} x_{\pi j}^T$ to get:

$$E_{\Pi_Y} \tilde{X}_i(t) \tilde{X}_j^T(t) - y_{\pi i} y_{\pi j}^T = E_{\Pi_Y} \int_0^t [a_{ij}(\tilde{X}(s)) + b_i(\tilde{X}(s)) x_j^T(s) + \tilde{X}_i(s) b_j^T(\tilde{X}(s))] ds,$$

$$E_y \tilde{X}_{\pi i}(t) \tilde{X}_{\pi j}^T(t) - y_{\pi i} y_{\pi j}^T = E_y \int_0^t [a_{\pi i \pi j}(\tilde{X}(s)) + b_{\pi i}(\tilde{X}(s)) \tilde{X}_{\pi j}^T(s) + \tilde{X}_{\pi i}(s) b_{\pi j}^T(\tilde{X}(s))] ds.$$

Thus, just as in (i) we get

$$a_{ij}(\Pi y) + b_i(\Pi y) y_{\pi j}^T + y_{\pi i} b_j^T(\Pi y) = a_{\pi i \pi j}(y) + b_{\pi i}(y) y_{\pi j}^T + y_{\pi i} b_{\pi j}^T(y),$$

so by (i) we get: $a_{ij}(\Pi y) = a_{\pi i \pi j}(y)$.

- (Sufficiency) For $g := f \circ \pi$, the chain rule gives

$$\nabla_{\pi i} g(x) = \nabla_i f(\Pi x), \text{ for all } \Pi \in Q(m) \text{ and } x \in S^m.$$

Hence,

$$\begin{aligned} Lg(x) &= \frac{1}{2} \sum_{i,j=1}^m (a_{\pi i \pi j}(x) \nabla_{\pi i})^T \nabla_{\pi j} g(x) + \sum_{i=1}^m b_{\pi i}^T(x) \nabla_{\pi i} g(x) \\ &= \frac{1}{2} \sum_{i,j=1}^m (a_{\pi i \pi j}(x) \nabla_i)^T \nabla_j f(\Pi x) + \sum_{i=1}^m b_{\pi i}^T(x) \nabla_i f(\Pi x) \\ &= Lf(\Pi x) \text{ by (2.3.4).} \end{aligned}$$

So $f(\tilde{X}(t)) - \int_0^t Lf(\tilde{X}(s)) ds = g(\tilde{X}(t)) - \int_0^t Lg(\tilde{X}(s)) ds$ is a

P_y -martingale, which implies that $f(\tilde{X}(t)) - \int_0^t Lf(\tilde{X}(s)) ds$ is a

$P_y \circ \tilde{\Pi}^{-1}$ -martingale. But $P_y \circ \tilde{\Pi}^{-1}(\tilde{X}(0)) = \Pi y = P_y(\tilde{X}(0)) = \Pi y =$

$P_y(\tilde{X}(0)) = y = 1$; since the solution to the martingale problem for

$(L, \Pi y)$ is assumed to be unique, it follows that $P_{\Pi y} = P_y \circ \tilde{\Pi}^{-1}$,

for all $y \in S^m$ and all $\tilde{\Pi} \in \tilde{Q}(m)$. $\square$

Note that in the process of proving the above theorem, we actually rederived a particular case of Proposition (2.1.5): namely,

that 1-fold exchangeability plus uniqueness of solution to the martingale problem implies exchangeability of the coordinates of the canonical process $\tilde{X}$ as random elements of $C([0, \infty); S)$.

Our next result singles out conditions on the generator L to ensure a unique solution to all martingale problems for the subsystem of our cloud of particles, yielding the Markovian behavior at all levels, as described at the beginning of this section.

2.3.8 Theorem: Let $\underline{a}$ and $\underline{b}$ satisfy the conditions of theorem (2.3.3). A necessary and sufficient condition for all the martingale problems associated with $L_J := \frac{1}{2} \sum_{i,j \in J} (a_{ij} \nabla_i)^T \nabla_j + \sum_{i \in J} b_i^T \nabla_i$, for each non-empty $J \subset \{1, 2, \dots, m\}$, to have a unique solution is that the following hold: there exist functions $\hat{b}_i : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $i=1, 2, \dots, m$; $\hat{a}_{ii} : \mathbb{R}^d \rightarrow$ the space of $d \times d$ matrices, $i=1, 2, \dots, m$; $\hat{a}_{ij} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow$ the space of $d \times d$ matrices such that:

$$2.3.9 \quad \begin{cases} b_i(x) = \hat{b}_i(x_i), \text{ for every } x \in \mathbb{R}^n, i=1, 2, \dots, m; \\ a_{ij}(x) = \begin{cases} \hat{a}_{ii}(x_i) & \text{when } i = j; \\ \hat{a}_{ij}(x_i, x_j) & \text{when } i \neq j; \end{cases} \text{ for every } x \in \mathbb{R}^n, \\ i, j = 1, 2, \dots, m. \end{cases}$$

Proof: (Sufficiency) Immediate from theorem (2.3.1).

(Necessity) Let P_x, P_z be the solutions to the martingale problems for L started respectively at x and z , i.e.

$$P_x \circ X(0) = \delta_x \quad \text{and} \quad P_z \circ \tilde{X}(0) = \delta_z.$$

Then we have

$$E_{x,1} \tilde{X}(t) - x_1 = E_x \int_0^t b_i(X(s)) ds$$

and
$$E_{\mathbf{z}} \tilde{X}_i(t) - z_i = E_{\mathbf{z}} \int_0^t b_i(\tilde{X}(s)) ds.$$

Defining the mapping $\phi_i : D([0, \infty) : S^m) \rightarrow D([0, \infty) : S)$ by $\phi_i(\omega)(t) := \omega_i(t)$, $i=1, 2, \dots, m$, it is clear that $P_x \circ \phi_i^{-1}$ and $P_z \circ \phi_i^{-1}$ are solutions to the martingale problem for $L_{\{i\}}$ started respectively at x_i and z_i . If $x_i = z_i$, then by assumption, $P_x \circ \phi_i^{-1} = P_z \circ \phi_i^{-1}$ so the left-hand sides of the above equalities are equal. Therefore, so are the right-hand sides and, as in the proof of theorem (2.3.3), $b_i(x) = b_i(z)$.

A similar argument, using the functions $\phi_{i,j}(\omega)(t) := (\omega_i(t), \omega_j(t))$, yields $a_{ij}(x) = a_{ij}(z)$ as soon as $x_i = z_i$, $x_j = z_j$.

As a consequence of theorems (2.3.3) and (2.3.8), we can now characterize those diffusions exhibiting both exchangeability and the strong Markov property within each subsystem. More precisely,

2.3.10 Corollary: Let $\underline{a}$ and $\underline{b}$ satisfy the conditions of theorem (2.3.3). Then a necessary and sufficient condition for both (2.3.4) and (2.3.9) to hold is that the corresponding triple (b, σ, ρ) defined by (2.3.6) satisfies (2.3.7) independently of the value of μ_m , i.e.

$$2.3.11 \quad \begin{cases} b_i(y) = b(y_i), & i=1, 2, \dots, m; \\ a_{ii}(y) = \sigma(y_i), & i=1, 2, \dots, m; \\ a_{ij}(y) = \rho(y_i, y_j), & i, j=1, 2, \dots, m, i \neq j. \end{cases}$$

2.4. On the existence of solutions for some martingale problems on $C([0, \infty) : P(\mathbb{R}^d))$

The final result of this chapter will be a characterization of the limiting empirical process associated with a sequence of exchangeable finite-dimensional diffusions by way of the martingale problem approach. We first identify the generator of the empirical process

associated with the coordinates of a canonical finite-dimensional diffusion. Under some additional conditions, we derive the tightness of the sequence, thereby obtaining the existence of a solution to a family of martingale problems for measure-valued processes. For a subfamily of these problems, uniqueness is known to hold because of the existence of dual processes constructed by Dawson and Kurtz [DK].

Weak laws of large numbers for diffusions have been investigated by many authors (for instance, Dawson [DA1], Léonard [LE], Oelschläger [OE]; this list is by no means exhaustive). In [DA1], the sequence of processes considered is described by stochastic differential equations of the forms $dx_i^m(t) = a(x_i^m(t))dt + \sigma(x_i^m(t))dw_i(t) +$

$$\frac{1}{m} \sum_{\substack{j=1 \\ j \neq i}}^m v_{ij}(x_i^m(t) - x_j^m(t))dt, \quad \text{for } i = 1, 2, \dots, m. \quad \text{Under some Lipschitz}$$

conditions on a , σ and v_{ij} , the equation's strong solution is unique and the empirical process $t \rightarrow \frac{1}{m} \sum_{i=1}^m \delta_{x_i^m(t)}$ is shown to converge weakly to a deterministic measure-valued process characterized as the unique $P(\mathbb{R}^d)$ -valued solution to a non-linear evolution equation. In [OE], the sequence is described by: for $i = 1, 2, \dots, m$,

$$dx_i^m(t) = b(x_i^m(t), \frac{1}{m} \sum_{i=1}^m \delta_{x_i^m(t)})dt + \sigma(x_i^m(t), \frac{1}{m} \sum_{i=1}^m \delta_{x_i^m(t)})dw_i(t).$$

and the unique weak limit is here the law of the solution to an associated stochastic differential equation, hence again deterministic.

In [LE], the system is $dx_i^m(t) = (\frac{1}{m} \sum_{j=1}^m b(x_i^m(t), x_j^m(t)))dt +$

$$(\frac{1}{m} \sum_{j=1}^m \sigma(x_i^m(t), x_j^m(t)))dw_i(t) \quad \text{for } i = 1, 2, \dots, m, \text{ a particular case of}$$

[OE]. However, the convergence is proven to hold in a stronger sense

here, on a subspace of $P(C[0, \infty); \mathbb{R}^d)$.

Our main goal in this section is to prove that, under some mild conditions, the empirical processes of a sequence of exchangeable diffusions are tight and any limit point of the sequence is a solution to a martingale problem associated with some operator on a subspace of $C(P(\mathbb{R}^d))$, thereby proving that a large class of such problems possess a solution. In many cases, these solutions will not be deterministic. In order to do so, we shall adapt the martingale problem formulation for the construction of our sequence, which does not necessarily guarantee the existence of some kind of stochastic differential equation for which a strong solution could be defined, as in the above cases. The formulation is therefore weaker, but more limits are obtained and, for the purpose of weak convergence, it is still strong enough.

Here still $S = \mathbb{R}^d$. Let $\tilde{X}^m(t) = (\tilde{X}_i^m(t))_{i=1}^m$ denote the S^m -valued canonical process associated with the martingale problem for L_m described in theorem (2.3.3) where, for $f \in C_c^\infty(S^m)$, we denote for all $x \in S^m$:

$$\begin{aligned} L_m f(x) &:= \frac{1}{2} (\underline{a}(x) \nabla)^T \nabla f(x) + \underline{b}^T(x) \nabla f(x) \\ &= L_m^1 f(x) + L_m^2 f(x) + L_m^3 f(x); \end{aligned}$$

$$L_m^1 f(x) := \sum_{i=1}^m b^T(x_i, \mu_m) \nabla_i f(x);$$

$$L_m^2 f(x) := \frac{1}{2} \sum_{i=1}^m (\sigma(x_i, \mu_m) \nabla_i)^T \nabla_i f(x);$$

$$L_m^3 f(x) := \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^m (\rho(x_i, x_j, \mu_m) \nabla_i)^T \nabla_j f(x);$$

$$\mu_m = \frac{1}{m} \sum_{i=1}^m \delta_{x_i}.$$

We begin by describing the martingale problem associated with the process $\frac{1}{m} \sum_{i=1}^m \delta_{\tilde{x}_i^m(t)}$, for some operator L_m .

Let $A \subset C(P(S))$ denote the algebra of so-called polynomials on $P(S)$, generated by functions of the form

$$F_f(\mu) := \langle f, \mu^{xk} \rangle = \int_S f d\mu^{xk}$$

where $f \in C(S^k)$, for some $k \in \mathbb{Z}^+$, and μ^{xk} denotes the k -fold product of measure μ by itself (i.e. $\mu^{xk} \in P(S)^k$). Note that, for $f \in C(S^k)$,

$$F_f(\mu_m) = \frac{1}{m^k} \sum_{j_1=1}^m \dots \sum_{j_k=1}^m f(x_{j_1}, \dots, x_{j_k})$$

which can be viewed as an element in $C(S^m)$. Computing L_m therefore reduces to determining the action of L_m on $F_f(\mu_m)$ as a function of coordinates $(x_1, x_2, \dots, x_m)$. This may be done as follows. Put $L_{m,f}^i(\mu_m) := L_m^i F_f(\mu_m)$, $i=1,2,3$, for $f \in \bigcup_{k \in \mathbb{Z}^+} C_c^\infty(S^k)$.

$$L_{m,f}^1(\mu_m) = \sum_{i=1}^m b^T(x_i, \mu_m) \nabla_i F_f(\mu_m)$$

$$= \sum_{i=1}^m b^T(x_i, \mu_m) \left[\frac{1}{m^k} \sum_{\alpha=1}^k \sum_{j_\alpha \text{ free}} \nabla_\alpha f(x_{j_1}, \dots, x_{j_{\alpha-1}}, x_i, x_{j_{\alpha+1}}, \dots, x_{j_k}) \right],$$

for $f \in C_c^\infty(S^k)$ and where $\sum_{j_\alpha \text{ free}}$ denotes the sum over all indices

$j_1, j_2, \dots, j_k$, save j_α , each going from 1 to m . Thus

$$L_m^1 F_f(\mu_m) = \sum_{\alpha=1}^k \left[\frac{1}{m} \sum_{i=1}^m b^T(x_i, \mu_m) \int_{S^{k-1}} \nabla_\alpha f(y_1, \dots, y_{\alpha-1}, x_i, y_{\alpha+1}, \dots, y_k) \times \right. \\ \left. \times d\mu_m^{x_{k-1}}(*) \right]$$

where $(*) = (y_1, \dots, y_{\alpha-1}, y_{\alpha+1}, \dots, y_k)$. So

$$L_m^1 F_f(\mu_m) = F_{B_k f}(\mu_m) \quad \text{where} \quad B_k f(y) := \sum_{\alpha=1}^k b^T(y_\alpha, \mu_m) \nabla_\alpha f(y).$$

$$\begin{aligned} L_m^2 F_f(\mu_m) &= \frac{1}{2} \sum_{i=1}^m (\sigma(x_i, \mu_m) \nabla_i)^T \nabla_i F_f(\mu_m) \\ &= \frac{1}{2} \sum_{i=1}^m \left[\frac{1}{m^k} \sum_{\alpha=1}^k \sum_{j_\alpha \text{ free}} (\sigma(x_i, \mu_m) \nabla_\beta)^T \nabla_\alpha f(x_{j_1}, \dots, x_{j_{\alpha-1}}, x_i, x_{j_{\alpha+1}}, \dots, x_{j_k}) \right] \\ &\quad + \frac{1}{2} \sum_{i=1}^m \left[\frac{1}{m^k} \sum_{\alpha, \beta=1}^k \sum_{j_\alpha, j_\beta \text{ free}} (\sigma(x_i, \mu_m) \nabla_\beta)^T \nabla_\alpha f(\dots, x_{j_{\alpha-1}}, x_i, x_{j_{\alpha+1}}, \dots, x_{j_{\beta-1}}, x_i, x_{j_{\beta+1}}, \dots) \right] \\ &= F_{\Sigma_k f}(\mu_m) + \frac{1}{2m} \sum_{\alpha \neq \beta} \left(\frac{1}{m} \sum_{i=1}^m \int_{S^{k-2}} (\sigma(x_i, \mu_m) \nabla_\beta)^T \nabla_\alpha f(\dots, x_i, \dots) \right. \\ &\quad \left. d\mu_m^{x_{k-2}}(**) \right) \end{aligned}$$

where $(**) = (y_1, \dots, y_{\alpha-1}, y_{\alpha+1}, \dots, y_{\beta-1}, y_{\beta+1}, \dots, y_k)$.

$$= F_{\Sigma_k f}(\mu_m) + \frac{1}{2m} F_{\Xi_k f}(\mu_m)$$

where $\Sigma_k f(y) := \frac{1}{2} \sum_{\alpha=1}^k (\sigma(y_\alpha, \mu_m) \nabla_\alpha)^T \nabla_\alpha f(y)$, $y \in S^k$,

$$\text{and } \Xi_k f(y) := \sum_{\substack{\alpha, \beta=1 \\ \alpha \neq \beta}}^k (\sigma(y_\alpha, \mu_m) \nabla_\beta)^T \nabla_\alpha f(y_{\alpha\beta}),$$

where $y \in S^k$ and $y_{\alpha\beta} = (y_1, y_2, \dots, y_{\beta-1}, y_\alpha, y_{\beta+1}, \dots, y_k)$.

$$\begin{aligned} L_m^3 F_f(\mu_m) &= \frac{1}{2} \sum_{i \neq j} (\rho(x_i^*, x_j, \mu_m) \nabla_i)^T \nabla_j F_f(\mu_m) \\ &= \frac{1}{2} \sum_{i \neq j} \left[\frac{1}{m^k} \sum_{\alpha \neq \beta} \sum_{\substack{j_\alpha, j_\beta \\ \text{free}}} (\rho(x_i, x_j, \mu_m) \nabla_\beta)^T \nabla_\alpha f(\dots x_{j_{\alpha-1}}, x_i, \dots x_{j_{\beta-1}}, x_i, x_{j_{\beta+1}}, \dots) \right] \\ &= F_{\Gamma_k} f(\mu_m) - \frac{1}{2m} \sum_{\alpha \neq \beta} \left(\frac{1}{m} \sum_{i=1}^m \int_{S^{k-2}} (\rho(x_i, x_i, \mu_m) \nabla_\beta)^T \nabla_\alpha f(\dots x_i \dots x_i \dots) d\mu_m^{x_{k-2}}(**) \right) \\ &= F_{\Gamma_k} f(\mu_m) - \frac{1}{2m} F_{T_k} f(\mu_m) \end{aligned}$$

$$\text{where } \Gamma_k f(y) := \frac{1}{2} \sum_{\substack{\alpha, \beta=1 \\ \alpha \neq \beta}}^k (\rho(y_\alpha, y_\beta, \mu_m) \nabla_\beta)^T \nabla_\alpha f(y), \quad y \in S^k,$$

$$\text{and } T_k f(y) := \frac{1}{2} \sum_{\substack{\alpha, \beta=1 \\ \alpha \neq \beta}}^k (\rho(y_\alpha, y_\beta, \mu_m) \nabla_\beta)^T \nabla_\alpha f(y_{\alpha\beta}).$$

$$\text{So } L_m^3 F_f(\mu_m) = F_{(B_k + \Sigma_k + \Gamma_k)} f(\mu_m) + \frac{1}{2m} F_{(\Xi_k - T_k)} f(\mu_m).$$

Thus, if $P_y^m \in (C([0, \infty); S^m))$ denotes the solution to the martingale problem for L_m started at $y \in S^m$, then $P_y^m \circ \phi_m^{-1} \in P(\dot{C}([0, \infty); P(S)))$ is a solution to the martingale problem for L_m started at $\frac{1}{m} \sum_{i=1}^m \delta_{y_i}$, where $\phi_m : C([0, \infty); S^m) \rightarrow C([0, \infty); P(S))$

is defined by $\Phi_m(\omega)(t) := \frac{1}{m} \sum_{i=1}^m \delta_{\omega_i}(t)$.

Remembering that $S = \mathbb{R}^d$, denoted by $\bar{S}$ the one-point compactification of S and let W^d be the set of $d \times d$ matrices with real entries.

2.4.1 Theorem: Let $b : S \times P(S) \rightarrow S$, $\sigma : S \times P(S) \rightarrow W^d$ and $\rho : S \times S \times P(S) \rightarrow W^d$ be continuous functions such that the following hold:

(i) there exists a $K > 0$ such that, for every $x, y \in S$ and $\mu \in P(S)$,

$$|b(x, \mu)|^2 + |\sigma(x, \mu)|^2 \leq K(1 + |x|^2)$$

and

$$|\rho(x, y, \mu)| \leq K(1 + |x|^2 + |y|^2);$$

(ii) For each $m \in \mathbb{Z}^+$ and each $y^m \in S^m$, the matrix $\underline{a}^m(y^m) \in W^{md}$ defined by (2.3.7) is positive definite and symmetric;

(iii) For each $m \in \mathbb{Z}^+$, $\nu^m \in P(S^m)$ is such that the coordinates are exchangeable random variables and $\sup_m \int |x_1|^4 \nu_m(dx) < \infty$;

then $\{P_{\nu^m}^m \circ \phi_m^{-1} : m \in \mathbb{Z}^+\} \subset P(C([0, \infty); P(\bar{S})))$ is tight.

Proof: Using Varadhan's metric d_{VAR} on $P(\bar{S})$ and the modulus of continuity mappings on $C([0, \infty); P(\bar{S}))$ and $C([0, \infty); \mathbb{R})$, we can show that a family $\{Q_m : m \in \mathbb{Z}^+\} \subset P(C([0, \infty); P(\bar{S})))$ is tight if and only if the family $\{Q_m \circ \Psi_f^{-1} : m \in \mathbb{Z}^+\} \subset P(C([0, \infty); \mathbb{R}))$ is tight for each $f \in M$, the set defined in Lemma (1.1.2). Here $\Psi_f(\omega)(t) := \int f d\omega(t)$. Since $M \subset C_c^\infty(S)$ may be assumed without loss of generality,

the theorem will be proved if we show that the sequence of processes $\{\langle f, \tilde{\mu}_m(\cdot) \rangle : m \in \mathbb{Z}^+\}$ is tight whenever $f \in C_c^\infty(\mathbb{R}^d)$, where $\tilde{\mu}_m(t) := \frac{1}{m} \sum_{i=1}^m \delta_{\tilde{X}_i^m(t)}$, $\tilde{X}^m(\cdot)$ denoting the canonical process on $C([0, \infty) : S^m)$.

- First observe that $\langle f, \tilde{\mu}_m(\cdot) \rangle$ is a D-semimartingale, for each $m \in \mathbb{Z}^+$ (see definition (1.3.1)). Take $\mathcal{C}$ to be the set of real-valued smooth functions on $\mathbb{R}$. The choice of mapping L^m requires some additional work. To this end, let $\mu_m := \frac{1}{m} \sum_{i=1}^m \delta_{x_i}$ for $x \in S^m$ and view $\langle f, \mu_m \rangle$ as an element of $C_c^\infty(S^m)$ as before. Then, for $\phi \in \mathcal{C}$, a computation like that done at the beginning of this section yields:

$$\begin{aligned} L_m \phi(\langle f, \mu_m \rangle) &= \phi'(\langle f, \mu_m \rangle) \left[\frac{1}{m} \sum_{i=1}^m b^T(x_i, \mu_m) \nabla f(x_i) + \frac{1}{2m} \sum_{i=1}^m (\sigma(x_i, \mu_m) \nabla)^T \nabla f(x_i) \right] \\ &\quad + \phi''(\langle f, \mu_m \rangle) \left[\frac{1}{2m^2} \sum_{i=1}^m (\nabla f(x_i))^T \sigma(x_i, \mu_m) \nabla f(x_i) + \right. \\ &\quad \left. \frac{1}{2m^2} \sum_{\substack{i,j=1 \\ i \neq j}}^m (\nabla f(x_j))^T \rho(x_i, x_j, \mu_m) \nabla f(x_i) \right]. \end{aligned}$$

Hence, by the formulation of the martingale problem for L_m , we know that $\langle f, \mu_m(t) \rangle - \int_0^t \beta_m(s) ds$ is a martingale with

$$\begin{aligned} \beta_m(s) &:= \frac{1}{m} \sum_{i=1}^m b^T(\tilde{X}_i^m(s), \tilde{\mu}_m(s)) \nabla f(\tilde{X}_i^m(s)) + \\ &\quad \frac{1}{2m} \sum_{i=1}^m (\sigma(\tilde{X}_i^m(s), \tilde{\mu}_m(s)) \nabla)^T \nabla f(\tilde{X}_i^m(s)). \end{aligned}$$

Then, with definition (1.3.1) in mind, we put

$$\begin{aligned}\alpha_m(s) &:= L_m(\langle f, \tilde{\mu}_m(s) \rangle)^2 - 2\langle f, \tilde{\mu}_m(s) \rangle \beta_m(s) \\ &= \frac{1}{m^2} \langle f, \tilde{\mu}_m(s) \rangle \left[\sum_{i=1}^m (\nabla f(\tilde{X}_i^m(s)))^T \sigma(\tilde{X}_i^m(s), \tilde{\mu}_m(s)) \nabla f(\tilde{X}_i^m(s)) \right. \\ &\quad \left. + \sum_{\substack{i,j=1 \\ i \neq j}}^m (\nabla f(\tilde{X}_j^m(s)))^T \rho(\tilde{X}_i^m(s), \tilde{X}_j^m(s), \tilde{\mu}_m(s)) \nabla f(\tilde{X}_i^m(s)) \right].\end{aligned}$$

Define $L^m(\phi, x, t, \omega_m) := \phi'(x) \beta_m(t, \omega_m) + \phi''(x) \alpha_m(t, \omega_m)$, where $\omega_m \in C([0, \infty); S^m)$. All the conditions for $\langle f, \mu_m(\cdot) \rangle$ to form a sequence of semimartingales are readily verified with the above parameters L^m, α_m, β_m .

- We must now verify the conditions of theorem (1.3.2) in order to ensure tightness. Condition (i) holds trivially, since f is bounded. For condition (ii), assume that the constant K in hypothesis (i) of theorem (2.4.1) is chosen large enough so that

$$\sup_{x \in S} |f(x)| + \sup_{x \in S} |\nabla f(x)| + \sup_{x \in S} |\nabla \nabla^T f(x)| < K, \text{ which is possible}$$

since $f \in C_c^\infty(S)$. Using the fact that, for $a \in \mathbb{W}^d, x, y \in \mathbb{R}^d$, the euclidean norm satisfies $|x^T a y| \leq |a| |x y^T| \leq |a| |x| |y|$, we get:

$$\begin{aligned}|\beta_m(t, \omega_m)|^2 &\leq \left(\frac{1}{m} \sum_{i=1}^m [|b(\tilde{X}_i^m(t), \tilde{\mu}_m(t))| |\nabla f(\tilde{X}_i^m(t))| + \right. \\ &\quad \left. \frac{1}{2} |\sigma(\tilde{X}_i^m(t), \tilde{\mu}_m(t))| |\nabla \nabla^T f(\tilde{X}_i^m(t))|] \right)^2 \\ &\leq 2K^2 \left(\frac{1}{m} \sum_{i=1}^m |b(\tilde{X}_i^m(t), \tilde{\mu}_m(t))|^2 + \frac{1}{4m} \sum_{i=1}^m |\sigma(\tilde{X}_i^m(t), \tilde{\mu}_m(t))|^2 \right)\end{aligned}$$

$$\begin{aligned}
& \leq 2K^3 \left(1 + \frac{1}{m} \sum_{i=1}^m |\tilde{X}_i^m(t)|^2\right) \\
|\dot{X}_m(t, \omega_m)| & \leq \frac{K}{m^2} \left\{ \sum_{i=1}^m |c(\tilde{X}_i^m(t), \tilde{\mu}_m(t))| |\nabla f(\tilde{X}_i^m(t))|^2 + \right. \\
& \quad \left. \sum_{\substack{i,j=1 \\ i \neq j}}^m |c(\tilde{X}_i^m(t), \tilde{X}_j^m(t), \tilde{\mu}_m(t))| |\nabla f(\tilde{X}_i^m(t))| |\nabla f(\tilde{X}_j^m(t))| \right\} \\
& \leq \frac{K^{7/2}}{m^2} \sum_{i=1}^m (1 + |\tilde{X}_i^m(t)|^2)^{1/2} + \frac{K^4}{m^2} \sum_{\substack{i,j=1 \\ i \neq j}}^m (1 + |\tilde{X}_i^m(t)|^2 + |\tilde{X}_j^m(t)|^2) \\
& \leq (K^{7/2} + K^4) \left(1 + \frac{2}{m} \sum_{i=1}^m |\tilde{X}_i^m(t)|^2\right).
\end{aligned}$$

We may thus select $C_t^m := 1 + \frac{1}{m} \sum_{i=1}^m |\tilde{X}_i^m(t)|^2$ in theorem (1.3.2), and check that condition (b) holds. It suffices to show that, for all $T > 0$,

$$2.4.2 \quad \sup_m E \sup_{t \leq T} |\tilde{X}_1^m(t)|^2 < \infty,$$

since our assumptions on the coefficients guarantee the exchangeability of $\{\tilde{X}_i^m(\cdot) : i=1, 2, \dots, m\}$ as a sequence of $C([0, \infty); S)$ -valued random variables. This implies that $\{\sup_{t \leq T} |\tilde{X}_i^m(t)|^2 : i=1, 2, \dots, m\}$ will be exchangeable as well. Condition (b) is then readily verified.

- Verification of (2.4.2).

$$R_t^m := |\tilde{X}_1^m(t)|^2 - \int_0^t (\text{trace } \sigma(\tilde{X}_1^m(s), \tilde{\mu}_m(s)) + 2(\tilde{X}_1^m(s))^T b(\tilde{X}_1^m(s), \tilde{\mu}_m(s))) ds$$

is a martingale. Let $V_T^m = \sup_{t \in [0, T]} |\tilde{X}_1^m(t)|^2$ then

$$V_T^m \leq \sup_{t \in [0, T]} |R_t^m| + \int_0^T 2K(1+V_s^m) ds \text{ by conditions (i) of (2.4.1)}$$

$$\leq 1 + \sup_{t \in [0, T]} |R_t^m|^2 + \int_0^T 2K(1+V_s^m) ds$$

$$\text{So, } EV_T^m \leq 1 + 4E|R_T^m|^2 + \int_0^T 2K(1+EV_s^m) ds$$

by Doob's martingale inequality (see Ethier and Kurtz, p. 64).

A Gronwall inequality (as in corollary (2.3.2)) yields

$$EV_T^m \leq (1 + 2KT + 4E|R_T^m|^2) e^{2KT}.$$

$$\text{Hence } \sup_m EV_T^m \leq (1 + 2KT + 4 \sup_m E|R_T^m|^2) e^{2KT}.$$

$$\text{Since } |R_t^m|^2 \leq 2(|\tilde{X}_1^m(t)|^4 + (\int_0^t d|\sigma(\tilde{X}_1^m(s), \mu_m(s))| + 2|(\tilde{X}_1^m(s))^T b(\tilde{X}_1^m(s), \mu_m(s))| ds)^2)$$

$$\leq 2(|\tilde{X}_1^m(t)|^4 + (\int_0^t 2dK(1+|\tilde{X}_1^m(s)|^2) ds)^2),$$

it suffices to prove that $\sup_m \sup_{0 \leq t \leq T} E|\tilde{X}_1^m(t)|^4 < \infty$.

Applying L_m to $f(x) := |x_1|^4$, where $x = (x_1, x_2, \dots, x_m)$, S^m yields

$$L_m f(x) = 2|x_1|^2 \text{trace}(\sigma(x_1, \mu_m)) + 4x_1^T \sigma(x_1, \mu_m) x_1 + 4|x_1|^2 b^T(x_1, \mu_m) x_1$$

$$|L_m f(x)| \leq (2d^2 + 4 + 4)K|x_1|^2(1 + |x_1|^2)$$

by the conditions on σ and b in (2.4.1)(i). Hence, by the martingale problem formulation, with $\tau_h := \inf\{t > 0 : |\tilde{X}_1^m(t)|^2 + |\tilde{X}_1^m(t)|^4 \geq h\}$, we get, for some constant C_1 independent of m and t ,

$$E|\tilde{X}_1^m(t \wedge \tau_h)|^4 \leq E|\tilde{X}_1^m(0)|^4 + \int_0^{t \wedge \tau_h} C_1(|\tilde{X}_1^m(s)|^2 + |\tilde{X}_1^m(s)|^4) ds.$$

Similarly, for $f(x) := |x_1|^2$ we get, for some C_2 independent of m and t ,

$$E|X_1^m(t \wedge \tau_h)|^2 \leq E|\tilde{X}_1^m(0)|^2 + \int_0^{t \wedge \tau_h} C_2(1 + |X_1^m(s)|^2) ds,$$

so by Grönwall's inequality applied to $1 + E|\tilde{X}_1^m(t \wedge \tau_h)|^2 + E|\tilde{X}_1^m(t \wedge \tau_h)|^4$ we obtain (as in corollary (2.3.2)):

$$1 + E|\tilde{X}_1^m(t)|^2 + E|\tilde{X}_1^m(t)|^4 \leq (1 + E|\tilde{X}_1^m(0)|^2 + E|\tilde{X}_1^m(0)|^4) e^{C_3 t},$$

for some C_3 independent of m and t .

This implies $\sup_m \sup_{0 \leq t \leq T} E|\tilde{X}_1^m(t)|^4 < \infty$ as soon as $\sup_m E|\tilde{X}_1^m(0)|^4 < \infty$. $\square$

At this point, one may wonder about condition (ii) of theorem (2.4.1): $\underline{a}^m$ needs to be positive definite everywhere for all values of m . Are there cases where this condition may be readily demonstrated? The following corollary answers affirmatively.

2.4.3 Corollary: Let $b : P(S) \rightarrow S$; $\sigma, \rho : P(S) \rightarrow W^d$ be bounded continuous functions such that σ and ρ satisfy the following conditions:

- (i) $\sigma(\mu), \rho(\mu)$ are symmetric for every $\mu \in P(S)$;
- (ii) Both $\rho(\mu)$ and $\sigma(\mu) - \rho(\mu)$ are positive-definite for every $\mu \in P(S)$;
- (iii) Condition (iii) of theorem (2.4.1);

then $\{P_m^m \circ \phi_m^{-1} : m \in \mathbb{Z}^+\} \subset P(C([0, \infty); P(\bar{S})))$ is tight.

Proof: It suffices to prove that, if $\underline{a} \in W^{md}$ is a block matrix of the form $\underline{a} = (a_{ij})_{i,j=1,2,\dots,m}$, where $a_{ij} \in W^d$ is such that

$$a_{ij} = \begin{cases} \sigma & \text{if } i = j \\ \rho & \text{if } i \neq j \end{cases}$$

for some $\sigma, \rho \in W^d$, then $\underline{a}$ is positive definite if and only if both $(\sigma - \rho)$ and $(\sigma + (m-1)\rho)$ are positive definite. This elementary fact is a consequence of formula

$$\det(\underline{a} - \lambda I_{md}) = (\det(\sigma - \rho - \lambda I_d))^{m-1} \det(\sigma + (m-1)\rho - \lambda I_d).$$

2.4.4 Theorem: Assume the conditions of theorem (2.4.1) hold. Then, for any $\mu \in P(S)$ with $\int |x|^4 d\mu(x) < \infty$, there is a solution to the martingale problem for L_∞ , started at μ , with L_∞ defined as follows:

$$L_\infty F_f(\mu) := F_{A_k(\mu)} f(\mu), \text{ where for } f \in C_c^\infty(S^k), k \in \mathbb{Z}^+,$$

$$\begin{aligned} A_k(\mu) f(y) := & \frac{1}{2} \sum_{\alpha=1}^k (\sigma(y_\alpha, \mu) \nabla_\alpha)^T \nabla_\alpha f(y) + \\ & \frac{1}{2} \sum_{\substack{\alpha, \beta=1 \\ \alpha \neq \beta}}^k (\rho(y_\alpha, y_\beta, \mu) \nabla_\beta)^T \nabla_\alpha f(y) + \\ & \sum_{\alpha=1}^k b^T(y_\alpha, \mu) \nabla_\alpha f(y). \end{aligned}$$

Proof: The mappings $\omega \rightarrow \xi_m(\omega, t) := F_f(\omega(t)) - \int_0^t L_m F_f(\omega(s)) ds$

and

$$\omega \rightarrow \xi(\omega, t) := F_f(\omega(t)) - \int_0^t L_\infty F_f(\omega(s)) ds$$

are bounded and continuous for each fixed $t \geq 0$, since f has

compact support and the mappings b, σ, ρ are all continuous. Hence,

$$\|\xi_m(\cdot, t) - \xi(\cdot, t)\| \leq \frac{Ct}{m} \text{ for some } C > 0, \text{ so that for a subsequence}$$

$$P_m^m \circ \phi_m^{-1} \Rightarrow P_\mu \in P(C([0, \infty); P(\bar{S}))) \text{ we have}$$

$$\lim_{m \rightarrow \infty} \int_B \xi_m(\cdot, t) dP_m^m \circ \phi_m^{-1}(\cdot) = \int_B \xi(\cdot, t) dP_\mu(\cdot)$$

for every $B \in \mathcal{F}_t$, the natural filtration (at t) on $C([0, \infty); P(\bar{S}))$.

Since ξ_m is a $P_m^m \circ \phi_m^{-1}$ -martingale for each $m \in \mathbb{Z}^+$, then ξ is a P_μ -martingale as well.

To complete the proof, we must show that $P_\mu \in P(C([0, \infty); P(S)))$. The argument runs along the lines of corollary (2.3.2). Let $Z(t) := \int |x|^2 d\tilde{\mu}_t(x)$ where $\tilde{\mu}_t$ denotes the canonical process associated with P_μ . Let $\tau_h := \inf \{t > 0 : Z(t) \geq h\}$. Then

$$\begin{aligned} E_\mu Z(t \wedge \tau_h) - E_\mu Z(0) &= E_\mu \int_0^{t \wedge \tau_h} ds \int_S [\text{trace}(\sigma(y, \tilde{\mu}_s)) + 2y^T b(y, \tilde{\mu}_s)] d\tilde{\mu}_s(y) \\ &\leq \int_0^t 2dK E_\mu (1 + Z(s \wedge \tau_h)) ds. \end{aligned}$$

$$\text{So } 1 + E_\mu Z(t) \leq (1 + E_\mu Z(0)) e^{2dKt} < \infty. \quad \square$$

For completeness, we end by mentioning a result of Dawson and Kurtz [DK] on the uniqueness of solution for the martingale problem mentioned in theorem (2.4.4). Using the method of duality, they have obtained uniqueness in a large number of such problems for measure-valued processes. As a consequence of their theorem 4.4, we get a unique solution when the coefficients do not depend on μ .

2.4.5 Theorem: Let $b : S \rightarrow S$, $\sigma : S \rightarrow W^d$, $\rho : S \times S \rightarrow W^d$ be continuous functions satisfying the conditions of theorem (2.4.1). Then the solution to the martingale problem for (L, μ) is unique, whenever $\int |x|^4 d\mu(x) < \infty$.

Proof: see Dawson and Kurtz [DK]. $\square$

Finally, in [DA1], [LE] and [OE], the corresponding (well-posed) martingale problems all belong to the class above, with $\rho = 0$. The solution is deterministic, in each case a new instance of the McKean-Vlasov limit for large systems of interacting particles for which the correlation between two randomly chosen particles decreases to zero as the size of the system goes to infinity. In that case, the limit is completely characterized because the associated nonlinear McKean-Vlasov equation possesses a unique solution.

In the case where $\rho \neq 0$, theorem (2.4.4) yields the existence of "random McKean-Vlasov" limits. A system of stochastic differential equations corresponding to the martingale problem of theorem (2.4.4) has yet to be defined and solved in the case where $\rho \neq 0$.

CHAPTER 3

INTERACTING FLEMING-VIOT PROCESSES

In this chapter, we present an extension of the Ohta-Kimura model for the evolution of certain characteristics of individuals belonging to several interacting populations with slightly overlapping territories. The original model was proposed by Ohta and Kimura to describe the evolution of the distribution of allelic types in a large monoecious population where random genetic drift and mutation are occurring. We then construct a measure-valued process to model the limiting, macroscopic, long-term behavior of these interacting Ohta-Kimura processes, in the sense of Fleming and Viot made precise below; and characterize this limit using the martingale problem formulation of Stroock and Varadhan. Finally, we study certain features of this new process.

3.1 Function-Valued Duals for Interacting Measure-Valued Processes

Throughout this chapter, $S = \mathbb{R}^d$ and $\bar{S}$ its one-point compactification. Letting Π_r^n denote the set of all mappings from $\{1, 2, \dots, n\}$ into $\{1, 2, \dots, r\}$ and $A^r := \bigcup_{m \geq 1} (C_c^\infty(S^m) \times \Pi_r^m)$, we define $F(\underline{\mu}, (f, \pi))$ for each $\underline{\mu} = (\mu_1, \mu_2, \dots, \mu_r) \in P(\bar{S})^r$ and each $(f, \pi) \in A^r$, as follows:

$$3.1.1 \quad F(\underline{\mu}, (f, \pi)) := \int_{S^m} f \, d^m x \sum_{i=1}^m \mu_{\pi(i)}.$$

Occasionally, we write $F_{\underline{\mu}}(f, \pi) \stackrel{\text{def}}{=} F_{(f, \pi)}(\underline{\mu})$, for convenience.

By the Stone-Weierstrass theorem, the algebra generated by the constant functions on $\bar{S}$, together with all functions of the form $F(\cdot, (f, \pi))$ for some $(f, \pi) \in A^r$, is dense in $C(P(\bar{S})^r)$. This algebra possesses the additional feature of containing only smooth functions, in the following sense.

If we define the partial derivative of $F(\underline{\mu}) = F_{(f, \pi)}(\underline{\mu})$ with respect to μ_j in direction $x \in S$ by

$$\frac{\partial F(\underline{\mu})}{\partial \mu_j(x)} := \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [F((\mu_1, \mu_2, \dots, \mu_{j-1}, \mu_j + \epsilon x, \mu_{j+1}, \dots, \mu_r)) - F(\underline{\mu})],$$

then a computation yields, for any $x, y \in S$ and $j, h \in \{1, 2, \dots, r\}$,

$$\frac{\partial F(\underline{\mu})}{\partial \mu_j(x)} = \sum_{k \in \pi^{-1}(j)} \int_{S^{m-1}} f(x_1, \dots, x_{k-1}, x, x_{k+1}, \dots, x_m) d \prod_{\substack{i=1 \\ i \neq k}}^m \mu_{\pi(i)}(x_i)$$

and so

$$\frac{\partial^2 F(\underline{\mu})}{\partial \mu_j(x) \partial \mu_h(y)} = \sum_{\substack{k \in \pi^{-1}(j) \\ \ell \in \pi^{-1}(h) \\ k \neq \ell}} \int_{S^{m-2}} f(x_1, \dots, x_{k-1}, x, x_{k+1}, \dots, x_{\ell-1}, y, x_{\ell+1}, \dots, x_m) d \prod_{\substack{i=1 \\ i \neq k, \ell}}^m \mu_{\pi(i)}(x_i).$$

By analogy with the finite-dimensional case, many $P(S)^r$ -valued processes may be characterized as solutions to martingale problems with respect to second order partial differential operators. We shall only deal in this chapter with a special case, actually involving a difference operator as well, but several extensions will readily come to mind.

Note that $\{F(\cdot, (f, \pi)) : (f, \pi) \in A^r\}$ is itself an algebra. If,

for $(f_1, \pi_1), (f_2, \pi_2) \in A^{\mathbb{R}}$, we define $(f_1, \pi_1) \cdot (f_2, \pi_2) := (f, \pi)$,

where $f(x_1, \dots, x_{m_1+m_2}) := f_1(x_1, \dots, x_{m_1}) f_2(x_{m_1+1}, \dots, x_{m_1+m_2})$ and

$$\pi(m) = \begin{cases} \pi_1(m) & \text{if } m = 1, 2, \dots, m_1 \\ \pi_2(m-m_1) & \text{if } m = m_1+1, \dots, m_1+m_2, \end{cases} \quad \text{then clearly}$$

$$F(\mu, (f_1, \pi_1)) \cdot F(\mu, (f_2, \pi_2)) = F(\mu, (f_1, \pi_1) \cdot (f_2, \pi_2)).$$

Similarly, for sums, $F(\mu, (f_1, \pi_1)) + F(\mu, (f_2, \pi_2)) = F(\mu, (f_1, \pi_1) + (f_2, \pi_2))$

with the appropriate definition of sum on $A^{\mathbb{R}}$. The product formula

will allow us to extend the duality equation to all finite-dimensional

distributions in the case of measure-valued processes. Recalling

definition (2.2.2), the $P(S)^{\mathbb{R}}$ -valued process X started at μ and

the $A^{\mathbb{R}}$ -valued process Y started at y are dual with respect to

F if, for every $t \geq 0$ we have

$$E_{\mu} F(X_t, y) = E_y F(\mu, Y_t).$$

Assuming that X and Y are time-homogeneous Markov processes

satisfying the conditions of definition (2.2.2), and y^1, y^2 are

independent copies of Y , started respectively at $y^1 \in A^{r_1}$ and

with the distribution of $y_{t-s}^1 \cdot y^2$, for $y^2 \in A^{r_2}$ and $t > s$,

we get

$$E_{\mu} (F(X_t, y^1) | F_s^X) = E_{y^1} F(X_s, y_{t-s}^1).$$

Hence, for any $0 \leq s < t < \infty$,

$$E_{\mu} F(X_t, y^1) F(X_s, y^2) = E_{\mu} [E_{y^1} F(X_s, y_{t-s}^1) F(X_s, y^2)]$$

$$\begin{aligned}
&= E_{\mu} E_{y^1} F(X_s, y_{t-s}^1 \cdot y^2) \\
&= E_{y^1} E_{\mu} F(X_s, y_{t-s}^1 \cdot y^2) \\
&= E_{y^1} E_{y_{t-s}^1 \cdot y^2} F(\mu, y_s^2) .
\end{aligned}$$

In general, for any choice of $0 < t_n < t_{n-1} < \dots < t_2 < t_1 < \infty$ and $y^1, y^2, \dots, y^n \in A^r$ for some r , with $y^1, y^2, \dots, y^n$ independent copies of Y , started respectively at y^1 , with the distribution of $y_{t_1-t_2}^1, y_{t_2-t_3}^2, \dots, y_{t_{n-1}-t_n}^{n-1}, y^n$, we get

$$3.1.2 \quad E_{\mu} \prod_{i=1}^n F(X_{t_i}, y^i) = E_{y^1} E_{y_{t_1-t_2}^1} \cdot y^2 \dots E_{y_{t_{n-1}-t_n}^{n-1}} \cdot y^n F(\mu, y_{t_n}^n) .$$

The finite-dimensional distributions of a $P(S)^r$ -valued process are thus expressed explicitly in terms of those of its A^r -valued dual, if the latter is a time-homogeneous Markov process, a fact that will be used in the proof of the scaling theorem.

3.1.3 Example: Let $\mu(t) = (\delta_{\beta_1(t)}, \delta_{\beta_2(t)}, \dots, \delta_{\beta_r(t)})$ where $\beta_1, \beta_2, \dots, \beta_r$ are independent standard Wiener processes (started at $0, S$). If H_t^m denotes the heat semigroup on $C(S^m)$, i.e.

$$H_t^m g(y) := \int_{S^m} g(x) (2\pi t)^{-md/2} \exp\{-|x-y|^2/2t\} dx ,$$

and, for $\pi \in \Pi_r^m$, $D_\pi : S^r \longrightarrow S^m$ is defined by

$$D_\pi(y_1, y_2, \dots, y_r) := (y_{\pi(1)}, y_{\pi(2)}, \dots, y_{\pi(m)}) ,$$

then for $(f, \pi) \in A^r$ with $f \in C_c^\infty(S^m)$, we have

$$\begin{aligned}
EF(\underline{\mu}(t), (f, \pi)) &= Ef \circ D_{\pi}(\beta_1(t), \beta_2(t) \dots \beta_r(t)) \\
&= H_t^r(f \circ D_{\pi})(0) \\
&= F(\underline{\mu}(0), (H_t^r(f \circ D_{\pi}), \pi)).
\end{aligned}$$

The dual of $\underline{\mu}(t)$ is therefore a purely deterministic,

$U(C_0^{\infty}(S^m) \times \Pi_r^m)$ -valued process. Even in this simple case, formula (3.1.2) is tedious to write down, but, since it will be used later, we do so nevertheless. For $0 < t_n < t_{n-1} < \dots < t_2 < t_1 < \infty$ and $(f^i, \pi^i) \in A^r$, $i = 1, 2 \dots n$, we have

$$E \prod_{i=1}^n F(\underline{\mu}(t_i), (f^i, \pi^i)) = H_{t_n - t_{n-1}}^{nr} (\dots H_{t_2 - t_3}^{2r} ((f^2 \circ D_{\pi^2}) \cdot H_{t_1 - t_2}^r (f^1 \circ D_{\pi^1})) \dots)(0).$$

3.2 Interacting Ohta-Kimura Processes

Suppose that r large populations of n individuals each (rn individuals in all) are under observation, each individual possessing a set of d characteristics which together determine its type. We are interested in the distributions of types in the r populations, as they evolve according to the following rules: each individual generates a large number of offsprings and dies. However, due to the environment, every population is stable. We make the simplifying assumption that it is maintained at n individuals, so that a single offspring per parent survives, on average. The surviving one is equipped with new characteristics chosen with high probability according to the empirical distribution of its parent's population (at the time of death of the parent), but occasionally from the empirical distribution of one of the other populations, a consequence of some

territorial overlap between the populations. The surviving offspring will sometimes be a mutant, i.e. one of its characteristics may be slightly altered. All types are assumed equally viable, i.e. the model is selectively neutral.

The use of another population's empirical distribution may be motivated by assuming that the chances of survival of the various types, among that portion of the offsprings invading a new territory, are best represented by the actual distribution of the population dwelling on that territory.

Assuming that each characteristic takes its values in the set of integer multiples of some basic unit, we see that the state space of our process is a countable subset of $P(\underline{Z}^d)^r$, namely the set $\mathcal{D}$ of r -tuples of discrete measures of the form $\frac{1}{n} \sum_{i=1}^n \delta_{x_i}$, for $x_1, x_2, \dots, x_n \in \underline{Z}^d$.

Let us sketch a construction of the process. Since each individual possesses a unique offspring, we may identify each original member of each population and its line of descendents with a tagged particle in motion in $\underline{Z}^d$. On a rich enough probability space, we build the following mutually independent random variables:

- $\{e_k^{i,j} : i=1,2,\dots,n; j=1,2,\dots,r; k=1,2,3,\dots\}$ exponentially distributed with respective means γ_j^{-1} ;
- $\{f_k^{i,j} : i=1,2,\dots,n; j=1,2,\dots,r; k=1,2,3,\dots\}$ exponentially distributed with respective means η_j^{-1} ;
- $\{u_\ell : \ell=1,2,3,\dots\}$ uniformly distributed over $\{1,2,\dots,n\}$;

- $\{v_\ell : \ell = 1, 2, 3, \dots\}$ uniformly distributed over $\{x \in \mathbb{Z}^d : |x| = 1\}$;
- $\{z_\ell : \ell = 1, 2, 3, \dots\}$ uniformly distributed over $\{1, 2, \dots, r\}$;
- $\{b_\ell : \ell = 1, 2, 3, \dots\}$ Bernoulli with mean $(r\varepsilon)/(r-1)$, $\varepsilon < 1/2$.

Next, define $\tau_\ell^{i,j} = \sum_{k=1}^{\ell} e_k^{i,j}$; the time of birth of the ℓ^{th} descendant of individual i of population j , and $\omega_\ell^{i,j} = \sum_{k=1}^{\ell} f_k^{i,j}$, the time of the ℓ^{th} mutation occurring in the line of descendants of individual i of population j . Since the probability of observing infinitely many elements of $\{\tau_\ell^{i,j}, \omega_\ell^{i,j} : i=1, 2, \dots, j; j=1, 2, \dots, r; \ell=1, 2, 3, \dots\}$ in any interval of time is zero, we may order that set of random times and relabel them $0 = s_0 < s_1 < s_2 < \dots$. Each s_h equals either $\tau_\ell^{i,j}$ or $\omega_\ell^{i,j}$ for some value of i, j, ℓ , so we may write $i = i(h)$, $j = j(h)$, $\ell = \ell(h)$ for those particular values corresponding to s_h .

The process is then constructed iteratively, denoting $X_n(t) = ((X_n(t))_j)_{j=1}^r$ started at some value $X_n(0) \in \mathcal{D}$, as follows:

$$X_n(t) = X_n(s_h) \text{ for all } t \in [s_h, s_{h+1});$$

$$(X_n(s_h))_j = \begin{cases} (X_n(s_{h-1}))_j & \text{if } j \neq j(h), h=1, 2, 3, \dots; \\ (X_n(s_{h-1}))_j + \frac{1}{n}(\delta_y - \delta_x) & \text{if } j=j(h), h=1, 2, 3, \dots \end{cases}$$

where x denotes particle $i(h)$ from population $j(h)$ and y is chosen as follows:

- If $s_h = \omega_\ell^{i,j}$ then $y = x + v_\ell$;
- If $s_h = \tau_\ell^{i,j}$ and $b_\ell = 0$ then y is particle $u_{j(h)}$ from population $j(h)$;

- If $s_h = \tau_{\ell}^{i,j}$ and $b_{\ell} = 1$ then y is particle $u_{j(h)}$ from population z_{ℓ} .

The process $X_n(t)$ is a continuous-time Markov chain with bounded jump rate, so, by the theory of Markov semigroups, if we denote by $\bar{Z}^d$ the one-point compactification of $\underline{Z}^d$, then the canonical process on $D([0, \infty): (\bar{Z}^d)^r)$ with the same finite-dimensional distributions and X_n 's is a solution to a martingale problem with respect to an operator L_n described as follows (see Ethier and Kurtz [EK], pp. 162-172 for the full motivation of the last statements).

For any $x, y \in \bar{S}$ and any $P \in P(\bar{S})$, let $P^{xy} := P + \frac{1}{n}(\delta_y - \delta_x)$, a finite measure; then define, for any $F(\mu, (f, \pi))$ of the form (3.1.1),

$$3.2.1 \quad L_n^F(f, \pi)(\mu) := \sum_{j=1}^r \sum_{x, y \in \bar{Z}^d} g_k(x, y) [F_{(f, \pi)}(\mu_1, \dots, \mu_{j-1}, \mu_j^{xy}, \mu_{j+1}, \dots, \mu_r) - F_{(f, \pi)}(\mu)],$$

where $g_j(x, y) := \eta_j \mu_j(x) \theta_{xy} + \gamma_j \mu_j(x) [(1-\epsilon) \mu_j(y) + \frac{\epsilon}{r-1} \sum_{\substack{i=1 \\ i \neq j}}^r \mu_i(y)]$;

$$\theta_{xy} := \begin{cases} 1 & \text{if } |x-y| = 1 \\ -2d & \text{if } x = y \\ 0 & \text{otherwise} \end{cases};$$

$\eta_j > 0$ denotes the mutation rate in population j ;

$\gamma_j > 0$ denotes the sampling rate in population j ;

$\epsilon > 0$ denotes the interpopulation sampling parameter.

It is clear from the construction above (and confirmed by a glance at the generator L_n) that, whenever $x_n(0) \in D$, then $x_n(t) \in D$ for all $t \in [0, \infty)$. These starting points will be the only ones of interest to us in the sequel, so the compactification is only a technical device here.

3.2.2 Remark: In this particular model, we note that mutations are allowed only to nearest-neighbour characters and that interpopulation sampling is done on an equal probability basis. The above construction need only be slightly altered to yield a class of interacting Ohta-Kimura processes for which the interpopulation sampling is done according to a Markov chain on $\{1, 2, \dots, r\}$. The new processes solve the martingale problems for L_n^q , given by formula (3.2.1) with the more complex coefficients

$$g_j^q(x, y) := n_j \mu_j(x) \theta_{xy} + \gamma_j \mu_j(x) \sum_{k=1}^r q_{jk} \mu_k(y),$$

where $(q_{jk})_{j,k=1,2,\dots,r}$ is a Markov transition matrix.

3.3 Interacting Fleming-Viot Processes: Existence

In their paper [FV], Fleming and Viot established the existence and uniqueness of a measure-valued diffusion as the limit, in the sense of weak convergence of processes, of a sequence of rescaled Ohta-Kimura processes, for which the population size n increases to infinity, the mesh of the lattice of possible types shrinks at a rate proportional to $n^{-1/2}$ and the speed of the evolution is scaled up a factor n^2 to ensure convergence. The process thus obtained

may serve to model the wandering distribution of types in a large population whose characteristics of interest (for instance, measurements) change continuously with time.

We now proceed in a similar fashion with a sequence of interacting Ohta-Kimura processes. In addition to the above changes of scale, we must also decrease the rate of interpopulation sampling by a factor n^{-1} in order to guarantee the occurrence of weak limits. We begin with a few technical results.

$$\begin{aligned}
 3.3.1 \text{ Lemma: } F_{(f,\pi)}(\mu_1, \dots, \mu_{j-1}, \mu_j^{xy}, \mu_{j+1}, \dots, \mu_r) &= F_{(f,\pi)}(\mu) + \\
 &\frac{1}{n} \left[\frac{\partial F_{(f,\pi)}(\mu)}{\partial \mu_j(y)} - \frac{\partial F_{(f,\pi)}(\mu)}{\partial \mu_j(x)} \right] + \\
 &\frac{1}{2n^2} \left[\frac{\partial^2 F_{(f,\pi)}(\mu)}{\partial \mu_j(y) \partial \mu_j(y)} - \frac{\partial^2 F_{(f,\pi)}(\mu)}{\partial \mu_j(y) \partial \mu_j(x)} + \frac{\partial^2 F_{(f,\pi)}(\mu)}{\partial \mu_j(x) \partial \mu_j(x)} \right] + \\
 &\frac{R(f, \pi, \mu, x, y, j)}{n^3}
 \end{aligned}$$

where, for every choice of f, π, μ, x, y and j , $|R(f, \pi, \mu, x, y, j)| \leq 2^{2m} \|f\|$ when $f \in C_c^\infty(S^m)$.

Proof: For any ν_1, ν_2 finite measures on $(\bar{S}, \mathcal{B}(\bar{S}))$, $(\nu_1 + \nu_2)^{\otimes m} = \sum_{\phi \in \Pi_2^m} \left(\sum_{i=1}^m \nu_{\phi(i)} \right)$, so by separating the terms where $\nu_2 = \frac{1}{n}(\delta_y - \delta_x)$ occurs 0, 1, 2 or more than 2 times we get the above four terms, respectively. $\square$

Let us denote the rescaled version of X_n by X_n , i.e. define

$$(X_n(t))_i(A) := (X_n(n^2 t))_i(n^{-1/2}A) \quad \text{for all } A \in B(S),$$

with the interpopulation sampling rate ε replaced by εn^{-1} . The generator G_n of process X_n will now be derived.

3.3.2 Proposition: For all $(f, \pi) \in A^r$ and all $\mu \in P(\bar{S})^r$,

$$G_n F_{(f, \pi)}(\mu) = G_\infty F_{(f, \pi)}(\mu) + n^{-1/2} R(f), \quad \text{where}$$

$$G_\infty F_{(f, \pi)}(\mu) := \sum_{j=1}^r [\eta_j < \frac{\Delta \partial F_{(f, \pi)}(\mu)}{\partial \mu_j}, \mu_j > +$$

$$\gamma_j < \frac{\partial^2 F_{(f, \pi)}(\mu)}{\partial \mu_j \partial \mu_j}, Q(\mu_j) > +$$

$$\gamma_j \frac{\varepsilon}{r-1} \sum_{k=1}^r < \frac{\partial F_{(f, \pi)}(\mu)}{\partial \mu_j}, \mu_k - \mu_j >];$$

$R(f)$ is a bounded remainder term, independent of n ;

Δ denotes the Laplacian on $C_c^\infty(S^m)$;

$$Q(\mu_j)(x, y) := \mu_j(x) \delta_x(y) - \mu_j(x) \mu_j(y).$$

Proof: Remembering that the change in time scale $t \rightarrow n^2 t$ multiplies the generator by a factor n^2 , that we have made the change $\varepsilon \rightarrow n^{-1} \varepsilon$ and noting $u_x := x n^{-1/2}$, we have, from (3.2.1), that

$$G_n F_{(f, \pi)}(\mu) = n^2 \sum_{j=1}^r \sum_{x, y \in \bar{S}} g_j^n(x, y) [F_{(f, \pi)}(\mu_1, \dots, \mu_{j-1}, \mu_j^{u_x u_y}, \mu_{j+1}, \dots, \mu_r) - F_{(f, \pi)}(\mu)]$$

where

$$g_j^n(s, y) = \eta_j \mu_j(x) \theta_{xy} + \gamma_j \mu_j(x) (1 - \varepsilon/n) \mu_j(y) + \frac{\varepsilon}{n(r-1)} \sum_{\substack{i=1 \\ i \neq j}}^r \mu_i(y).$$

To simplify the expressions, let $F(v) := F_{\pi}(f, \pi)(\mu_1, \dots, \mu_{j-1}, v, \mu_{j+1}, \dots, \mu_r)$ where $(f, \pi) \in A^r$ and $j \in \{1, 2, \dots, r\}$ are fixed throughout the computation. $R(f)$ will denote some bounded terms, not necessarily the same throughout the calculations.

- The Mutation Terms: By Lemma (3.3.1), we have

$$\begin{aligned}
 & n^2 \sum_{x, y \in \underline{Z}} \mu_j(x) \theta_{xy} (F(\mu_j^{u_x u_y}) - F(\mu_j)) \\
 &= n \sum_{x, y \in \underline{Z}} \mu_j(x) \theta_{xy} \left[\frac{\partial F(\mu_j)}{\partial \mu_j(u_y)} - \frac{\partial F(\mu_j)}{\partial \mu_j(u_x)} \right] + \\
 & \quad \frac{1}{2} \sum_{x, y \in \underline{Z}} \mu_j(x) \theta_{xy} \left[\frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_y) \partial \mu_j(u_y)} - \frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_y) \partial \mu_j(u_x)} + \right. \\
 & \quad \left. \frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_x) \partial \mu_j(u_x)} \right] + n^{-1} R(f) \\
 &= n \sum_{x, y \in \underline{Z}} \mu_j(x) \theta_{xy} \left[\frac{\partial F(\mu_j)}{\partial \mu_j(u_y)} - \frac{\partial F(\mu_j)}{\partial \mu_j(u_x)} \right] + n^{-1} R(f) \text{ since } f \text{ is} \\
 & \quad \text{smooth with compact support.} \\
 &= \sum_{x, y \in \underline{Z}} \mu_j(x) \Delta \left(\frac{\partial F(\mu_j)}{\partial \mu_j(z)} \right) \Big|_{z=u_x} + n^{-1} R(f) = \left\langle \frac{\Delta \partial F(\mu_j)}{\partial \mu_j}, \mu_j \right\rangle + n^{-1} R(f)
 \end{aligned}$$

since $\sum_{y \in \underline{Z}} \theta_{xy} (f(u_y) - f(u_x)) = n^{-1} \Delta f(u_x) + n^{-2} R(f)$ whenever f is smooth with compact support.

- The Sampling Terms, Within the Same Population: By lemma (3.3.1),

$$n^2 \sum_{x, y \in \underline{Z}} \mu_j(x) \mu_j(y) (F(\mu_j^{u_x u_y}) - F(\mu_j))$$

$$\begin{aligned}
&= \frac{1}{2} \sum_{x, y \in \underline{Z}} \mu_j(x) \mu_j(y) \left[\frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_y) \partial \mu_j(u_y)} - \frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_y) \partial \mu_j(u_x)} + \right. \\
&\quad \left. \frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_x) \partial \mu_j(u_x)} \right] + \frac{R(f)}{n} \\
&= \sum_{x \in \underline{Z}} \mu_j(x) \frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_x) \partial \mu_j(u_x)} - \sum_{x, y \in \underline{Z}} \mu_j(x) \mu_j(y) \frac{\partial^2 F(\mu_j)}{\partial \mu_j(u_y) \partial \mu_j(u_x)} + \frac{R(f)}{n} \\
&= \left\langle \frac{\partial^2 F(\mu_j)}{\partial \mu_j \partial \mu_j}, Q(\mu_j) \right\rangle + n^{-1} R(f).
\end{aligned}$$

- The Sampling Terms, Between the Various Populations: By lemma (3.3.1),

$$\begin{aligned}
&n \cdot \sum_{x, y \in \underline{Z}} \mu_j(x) \mu_i(y) (F(\mu_j^{u_x u_y}) - F(\mu_j)) \\
&= \sum_{x, y \in \underline{Z}} \mu_j(x) \mu_i(y) \left[\frac{\partial F(\mu_j)}{\partial \mu_j(u_y)} - \frac{\partial F(\mu_j)}{\partial \mu_j(u_x)} \right] + n^{-1} R(f) \\
&= \left\langle \frac{\partial F(\mu_j)}{\partial \mu_j}, \mu_i - \mu_j \right\rangle + n^{-1} R(f).
\end{aligned}$$

The result now follows from the three equalities. $\square$

3.3.3 Remark: Just as in remark (3.2.2) above, if interpopulation sampling is performed following a Markov chain q instead of uniform selection, the generator of the rescaled interacting processes is the same, with expression " $\sum_{k=1}^r q_{jk}$ " instead of " $\frac{\epsilon}{r-1} \sum_{k=1}^r$ ". Here, of course, the rate of sampling must be rescaled appropriately, i.e.

$$q_{jk} \rightarrow \begin{cases} q_{jk}/n & \text{if } j \neq k \\ 1-n^{-1} \sum_{\substack{l=1 \\ l \neq j}}^r q_{jl} & \text{if } j = k. \end{cases}$$

We can now turn to the proof of the existence of interacting Fleming-Viot processes.

3.3.4 Theorem: The sequence of processes $\{\chi_n : n \in \mathbb{Z}^+\}$ is tight in $P(D([0, \infty) : P(\bar{S})^r))$.

Proof: It suffices to prove that the sequences of processes

$\{F_{(f, \pi)}(\chi_n) : n \in \mathbb{Z}^+\}$ are tight in $P(D([0, \infty) : \mathbb{R}))$ for each $(f, \pi) \in \Lambda^r$ (see Appendix for a proof). Because of Dynkin's lemma, we know

that $F_{(f, \pi)}(\chi_n(t)) - \int_0^t G_n F_{(f, \pi)}(\chi_n(s)) ds$ is a $\mathcal{B}_t$ -martingale, where

$\{\mathcal{B}_t\}$ is the natural filtration on $B(D([0, \infty) : P(\bar{S})^r))$, so $F_{(f, \pi)}(\chi_n)$

is a D-semimartingale for each $n \in \mathbb{Z}^+$: simply put $X^n = F_{(f, \pi)}(\chi_n)$,

$F_t^n = \mathcal{B}_t$, $C =$ the polynomials on $\mathbb{R}$, $A_t^n = t$ for every $t \in [0, \infty)$

and $L^n(\phi, x, t, \omega) := G_n \phi \circ F_{(f, \pi)}(\chi_n(t, \omega))$ in definition (1.3.1). It

suffices therefore to prove that the conditions of Joffe and Métivier

in theorem (1.3.2) are satisfied. Condition (iii) is clear while

condition (i) holds because $|F_{(f, \pi)}(\chi_n(0, \omega))| \leq \|f\|$ for all $n \in \mathbb{Z}^+$.

Next, there is a constant $C(f) > 0$ independent of $n \in \mathbb{Z}^+$ such that

(see theorem (1.3.2) for the notation)

$$|\beta_n(x, t, \omega)| = |G_n F_{(f, \pi)}(\chi_n(t, \omega))| \leq C(f) (\|\Delta f\| + \|f\|)$$

and

$$\begin{aligned} |\underline{\alpha}_{-n}(x, t, \omega)| &\leq |G_n(F_{(f, \pi)}(\chi_n(t, \omega)))^2| + 2|x| |\underline{\beta}_{-n}(x, t, \omega)| \\ &\leq C(f) (\|\Delta f\| \|\underline{f}\| + \|\underline{f}\|^2 + \|\underline{f}\|) + 2|x| |\underline{\beta}_{-n}(x, t, \omega)| \end{aligned}$$

so $|\underline{\beta}_{-n}(x, t, \omega)|^2 + |\underline{\alpha}_{-n}(x, t, \omega)| \leq K(f)(1+|x|^2)$ trivially for some constant $K(f)$ independent of $n \in \mathbb{Z}^+$, which implies condition (ii) with $C_t^n \equiv 1$! $\square$

3.3.5 Theorem: Given any $\underline{\mu} \in P(S)^r$, the martingale problem for (G_∞, δ_μ) , where G_∞ is the operator defined in proposition (3.3.2), possesses at least one solution. All solutions belong to $P(D([0, \infty): P(S)^r))$, so there is no explosion in finite time, under the additional condition that the second moments of μ_i are finite, for each $i = 1, 2, \dots, r$.

Proof: By lemma (A.1) of the appendix, the mappings $\phi_n : D([0, \infty): P(\bar{S})^r) \rightarrow D([0, \infty): \mathbb{R})$, defined by $\phi_n(\omega)(t) := F_{(f, \pi)}(\omega(t)) - \int_0^t G_n F_{(f, \pi)}(\omega(s)) ds$ for every $n \in \mathbb{Z}^+ \cup \{\infty\}$, are continuous, so for any subsequence of rescaled processes such that $P^n \circ \chi_n^{-1} \Rightarrow P$ for some $P \in P(D([0, \infty): (\bar{S})^r))$, we have that $P^n \circ \chi_n^{-1} \circ \phi_\infty^{-1} \Rightarrow P \circ \phi_\infty^{-1}$. This implies that $P^n \circ \chi_n^{-1} \circ \phi_\infty^{-1} \circ \tilde{X}_t \Rightarrow P \circ \phi_\infty^{-1} \circ \tilde{X}_t^{-1}$, where as before $\tilde{X}_t(\omega) := \omega(t)$, holds for all but countably many $t \in [0, \infty)$ by a well-known property of weak convergence on D (see [EK], p.131). But since

$$\begin{aligned} \|\tilde{X}_t \circ \phi_n - \tilde{X}_t \circ \phi_\infty\| &\leq \int_0^t |n^{-1/2} R(f)| ds \text{ by proposition (3.3.2)} \\ &\leq n^{-1/2} t |R(f)|, \end{aligned}$$

actually $P^n \circ \chi_n^{-1} \circ \phi_n^{-1} \circ \tilde{X}_t^{-1} \Rightarrow P \circ \phi_\infty^{-1} \circ \tilde{X}_t^{-1}$ for all but countably many $t \in [0, \infty)$. Now $\phi_n \circ \chi_n$ is a sequence of uniformly bounded right-continuous $\mathcal{B}_t$ -martingales, so by Lebesgue's dominated convergence theorem, for every $F \in \mathcal{B}_s$ and every $t \geq s \geq 0$,

$$\lim_{n \rightarrow \infty} \int_F \phi_n(\omega(t)) dP^n \circ \chi_n^{-1}(\omega) = \lim_{n \rightarrow \infty} \int_F \phi_n(\omega(s)) dP^n \circ \chi_n^{-1}(\omega)$$

and thus $\int_F \phi_\infty(\omega(t)) dP(\omega) = \int_F \phi_\infty(\omega(s)) dP(\omega)$ for all but countably many pairs (s, t) , hence by right-continuity of $\phi_\infty(\omega(\cdot))$, for all pairs (s, t) . This implies that P is a solution to the martingale problem for (G_∞, μ) , as long as $\chi_n(0)$ converges to $\mu \in P(S)^r$. This is always possible since the set $\mathcal{D}$ defined at the beginning of section (3.2) is dense in $P(S)^r$.

The proof that no explosion occurs in finite time runs along the lines of that in theorem (2.4.4), so we only sketch it. Denote the canonical process on $D([0, \infty): P(\bar{S})^r)$ by $\tilde{X}(t, \cdot) = (\tilde{X}_j(t, \cdot))$, $j = 1, 2, \dots, r$. Let $v(t) := \sum_{j=1}^r \int_S |x|^2 d\tilde{X}_j(t, x)$ and define $\tau_h := \inf\{t > 0: v(t) \geq h\}$. Then

$$\begin{aligned} E v(t \wedge \tau_h) - E v(0) &= \int_0^{t \wedge \tau_h} E G_\infty v(s) ds \\ &= \int_0^{t \wedge \tau_h} \left(2d \sum_{j=1}^r \eta_j + \sum_{j=1}^r \gamma_j \frac{\epsilon}{r-1} E(v(s) - \right. \\ &\quad \left. r \int_S |x|^2 d\tilde{X}_j(s, x)) \right) ds \\ &\leq \int_0^t (a + b E v(s \wedge \tau_h)) ds \quad \text{for some constants } a > 0 \end{aligned}$$

and $b > 0$, independently of h . Gronwall's inequality implies

$$Ev(t) \leq \left[\frac{a}{b} + Ev(0) \right] e^{bt} - \frac{a}{b} \quad \text{for all } t \geq 0. \quad \square$$

3.4 Interacting Fleming-Viot Processes: Uniqueness

We next prove that the martingale problem stated in theorem (3.3.5) possesses a unique solution by constructing a function-valued dual process, as done by Dawson and Hochberg in [DH2] for the original Fleming-Viot process. In view of theorem (2.2.4), we need the following

3.4.1 Lemma: Let G_∞ be the operator defined in proposition (3.3.2). Then

$$\begin{aligned} G_\infty F(f, \pi)(\mu) &= F(\mu, \left(\sum_{j=1}^m \eta_{\pi(j)} \Delta_j f + \sum_{\substack{j,k=1 \\ \pi(j)=\pi(k)}}^m \gamma_{\pi(j)} (\phi_{jk} f - f), \pi \right)) \\ &\quad + \sum_{j=1}^m \sum_{p=1}^r \gamma_{\pi(j)} \frac{\varepsilon}{r-1} (F(\mu, (f, \Xi_{jp} \pi)) - F(\mu, (f, \pi))), \end{aligned}$$

where $\Delta_j f$ denotes the Laplacian acting on the j^{th} coordinate of $f \in C_c^\infty(S^m)$; in short (see p. 71 for the exact definition),

$\phi_{jk} : C(S^m) \rightarrow C(S^{m-1})$ is defined, for $m > 1$ by

$$\phi_{jk} f(x_1, \dots, x_m) := f(x_1, \dots, x_{k-1}, x_j, x_{k+1}, \dots, x_m);$$

$\phi_{jk} f = f$ when $f \in C(S)$;

$$\Xi_{jp} : \Pi_r^m \rightarrow \Pi_r^m \text{ is defined by } \Xi_{jp}(\pi)(k) = \begin{cases} \pi(k) & \text{if } k \neq j \\ p & \text{if } k = j. \end{cases}$$

Proof: We only derive the second term, the first is done similarly.

Since

$$\left\langle \frac{\partial F(f, \pi)(\mu)}{\partial \mu_q}, \mu_p - \mu_q \right\rangle = \sum_{\substack{j=1 \\ \pi(j)=q}}^m (F(\mu, (f, \Xi_{jp}(\pi))) - F(\mu, (f, \pi))),$$

we get

$$\begin{aligned} & \sum_{q=1}^r \gamma_q \frac{\epsilon}{r-1} \sum_{p=1}^r \left\langle \frac{\partial F(f, \pi)(\mu)}{\partial \mu_q}, \mu_p - \mu_q \right\rangle \\ &= \sum_{j=1}^m \gamma_{\pi(j)} \frac{\epsilon}{r-1} \sum_{p=1}^r (F(\mu, (f, \Xi_{jp}(\pi))) - F(\mu, (f, \pi))). \quad || \end{aligned}$$

G_∞ may therefore be viewed as an operator on a subset of $C(A^r)$; it turns out to be the generator of a Markov process on A^r which we now construct. We only treat the case $r = 2$ for clarity, the general following in a similar way. We first recall the construction in [DH2], expressing it in terms of an operator-valued process R_t as it makes the description of our dual somewhat easier.

For any fixed positive integer m , let $t_i, i=1, 2, \dots, m$ be independent exponential random variables with respective means $E t_i = (\gamma i(i-1))^{-1}$, where $\gamma > 0$ represents the sampling rate. Denoting $s_{m-1} = t_m, s_{m-2} = t_m + t_{m-1}, \dots, s_1 = t_m + t_{m-1} + \dots + t_2$ (the successive times of jump of the process R_t being built), and $u_i, i=1, 2, \dots, m$ independent random variables distributed uniformly on the $i(i-1)$ pairs of distinct elements from $\{1, 2, \dots, i\}$, respectively; one defines recursively,

$$R_t := H_{2\gamma t}^m \quad \text{for } t \in [0, s_{m-1})$$

$$:= H_{2\eta(t-s_j)}^j \circ \phi_{u_{j+1}}^{j+1} \circ R_{s_j^-} \quad \text{for } t \in [s_j, s_{j+1}), j=1, 2, \dots, m-1,$$

with $s_0 = \infty$ and $\eta > 0$ representing the mutation rate. Here, for each $g \in C(S^m)$, $H_t^m g$ denotes the m -dimensional heat semigroup at g , i.e.

$$H_t^m g(x) := \begin{cases} g(x) & \text{at } t = 0 \\ \left(\frac{1}{\sqrt{2\pi t}} \right)^{dm} \int_{S^m} e^{-|x-y|^2/2t} g(y) dy & \text{for } t > 0; \end{cases}$$

and $\phi_{jk}^m : C(S^m) \rightarrow C(S^{m-1})$ is defined by

$$(\phi_{jk}^m f)(y_1, y_2, \dots, y_{m-1}) := f(x_1, \dots, x_m) \quad \text{for } j \neq k, \text{ where}$$

$$x_i = y_i \quad \text{for } i = 1, 2, \dots, k-1$$

$$x_k = y_j \quad \text{if } j < k$$

$$x_k = y_{j-1} \quad \text{if } j > k$$

$$x_i = y_{i-1} \quad \text{for } i = k+1, \dots, m.$$

Note that R_t is well-defined as a stochastic process taking its values in the space of bounded linear operators on $C(S^m)$.

Actually, $\|R_t g\| \leq \|g\|$, for all $t \geq 0$, and the strong continuity in t of H_t^m guarantees that $R_{s_j^-}$ exists.

More generally, for any subset A of the coordinates $\{x_1, x_2, \dots, x_m\}$, we may define R_t^A at $g \in C(S^m)$ by letting R_t act only on the coordinates of g that belong to A , leaving the other ones fixed for all times $t \geq 0$. Denoting by A^c the complement of A in $\{x_1, x_2, \dots, x_m\}$, we see that

$$R_t^A R_t^{A^C} = R_t^{A^C} R_t^A \text{ on } C(S^m), \text{ for all } t \geq 0 \text{ and all choices of } A.$$

We may now exhibit the dual process to our two-population process ($r=2$), noted $(f, \pi)_t = (f_t, \pi_t)$, started at $(f_0, \pi_0) \in C(S^m) \times \Pi_2^m$ for some fixed m . We use the fact that the process remains in that set forever, if one views $C(S^k) \times \Pi_2^m$ as a subset of $C(S^m) \times \Pi_2^m$ when $k < m$.

We first build recursively the process π_t , started at $\pi_0 \in \Pi_2^m$, as that process with independent coordinates $\pi_t(j)$, $j=1, 2, \dots, m$, each being a continuous-time Markov chain on $\{1, 2\}$ with exponential holding times $\tau_{j,k+1}$ with mean $E\tau_{j,k+1} = (\epsilon \gamma_{\pi_{\tau_{j,k}}(j)})^{-1}$, for $k = 0, 1, 2, \dots$ with $\tau_{j,0} = 0$.

Let $0 = \tau_0 < \tau_1 < \tau_2 < \dots$ denote the successive times of jump of process π_t . We may now build f_t , started at $f_0 \in C(S^m)$. After constructing a countable collection $\{(i,k) R_t^A : i=1, 2; k=0, 1, 2, \dots; A \in \{x_1, x_2, \dots, x_m\}\}$ of independent copies of process R_t^A , with the sampling rate γ of $(i,k) R_t^A$ equal to γ_i and the diffusion rate η equal to η_i , $i=1, 2$; define recursively

$$f_t := \sum_{k=0}^{\infty} I_{[\tau_k, \tau_{k+1})}(t) h_{t-\tau_k}^k, \text{ where}$$

$$h_0^0 := f_0, \quad h_0^k := h_{\tau_k}^{k-1} \text{ for } k = 1, 2, \dots$$

and

$$h_t^k := (2,k) R_t^{\pi_{\tau_k}^{-1}(2)} \circ (1,k) R_t^{\pi_{\tau_k}^{-1}(1)} (h_0^k), \quad k=0, 1, 2, \dots$$

We note that, after an almost surely finite time, $f_t \in C(S)$ and remains in that space from then on. Since $C(S^m)$ is separable

and the process $(f, \pi)_t$ is almost surely cadlag, there is a version of this process on $D([0, \infty): C(S^{\mathbb{N}}) \times \Pi_2^{\mathbb{N}})$: call it $(f, \pi)_t$. It is a solution to the martingale problem for G_{∞} on $C(A^2)$, started at $(f, \pi)_0$. Since $\|f_t\| \leq \|f_0\|$ for all $t \geq 0$, $|G_{\infty}F(\mu, (f, \pi)_s)| \leq C(f)$, a constant independent of $\varepsilon \leq 1$ and $s \geq 0$, which implies the integrability condition (iii) of theorem (2.2.4) and hence the duality equation, in this case written:

$$EF(X_0, (f, \pi)_t) = EF(X_t, (f, \pi)_0) \quad \text{for all } t \geq 0,$$

where X_t is any solution to the martingale problem of theorem (3.3.5), with values in $P(\bar{S})^{\mathbb{R}}$, and X and (f, π) are assumed independent.

3.4.2. Theorem: Given any $\mu \in P(S)^{\mathbb{R}}$, the martingale problem for $(G_{\infty}, \delta_{\mu})$ is well-posed (in $P(D([0, \infty): P(\bar{S})^{\mathbb{R}}))$). Further, if $\sum_{j=1}^{\mathbb{R}} \int |x|^2 d\mu_j(x) < \infty$, the solution belongs to $P(D([0, \infty): P(S)^{\mathbb{R}}))$.

Proof: This is a direct consequence of theorem (3.3.5), theorem (2.2.3) and the construction of a dual process. $\square$

3.4.3 Remark: Substituting " $q_{\pi(j)p}$ " for " $\varepsilon/(r-1)$ " in the formula for G_{∞} given in lemma (3.4.1) yields the dual form of the generalized operator mentioned in remarks (3.3.3) and (3.2.2). The construction of a dual given above may be modified to yield the characterization of these interacting Fleming-Viot processes (theorem (3.4.2)) as well.

3.5 The Scaling Theorem

Heuristically, the scaling theorem (3.5.1) below provides a picture of the macroscopic long-term behavior of a family of interacting Fleming-Viot processes.

Let $\tilde{Y}(t) := (Y_t^i)_{i=1,2,\dots,r}$ denote the canonical process associated with the solution to the martingale problem on $P(D([0,\infty); P(\bar{S})^r))$, for G_∞ defined in (3.3.2), started at $\tilde{Y}(0) \in P(\bar{S})^r$. Define

$$\tilde{Y}_n(t; A_1, A_2, \dots, A_r) := (Y_{n^{-1}t}^i(nA_i))_{i=1,2,\dots,r}$$

for all $t \in [0, \infty)$, $A_1, A_2, \dots, A_r \in \mathcal{B}(\bar{S})$ and $nA := \{x \in \bar{S} : xn^{-1} \in A\}$.

Then, if we rescale the parameter $\varepsilon = \varepsilon(n)$, the generator of $\tilde{Y}_n$ is given by

$$\begin{aligned} J_n F(f, \pi)(\mu) &:= F(\mu, (\sum_{j=1}^m \eta_{\pi(j)} \Delta_j f + n^2 \sum_{\substack{j,k=1 \\ \pi(j)=\pi(k)}}^m \gamma_{\pi(j)} (\phi_{jk} f - f), \pi)) \\ &+ \frac{n^2 \varepsilon(n)}{r-1} \sum_{j=1}^m \sum_{p=1}^r \gamma_{\pi(j)} (F(\mu, (f, \varepsilon_{jp} \pi)) - F(\mu, (f, \pi))). \end{aligned}$$

3.5.1 Theorem: Let $\beta_1, \beta_2, \dots, \beta_r$ denote independent standard Wiener processes in $\mathbb{R}^d$. Then $\lim_{n \rightarrow \infty} n^2 \varepsilon(n) = a \in [0, \infty)$ implies that the finite-dimensional distributions of processes $\tilde{Y}_n$ converge to those of process $\tilde{Y}_\infty$, where, for each $t \in [0, \infty)$, $\tilde{Y}_\infty(t) := (\delta_{\alpha_i(t)})_{i=1,2,\dots,r}$ with

- If $a = 0$ then $\alpha_i(t) := \beta_i(2\eta_i t)$, $i = 1, 2, \dots, r$;
- If $a \in (0, \infty)$ then define recursively, for each $i = 1, 2, \dots, r$,

$$\alpha_i(t) := \beta_i(2\eta_i t) - \beta_i(2\eta_i \tau_k^i) + \alpha_{u_k^i}(\tau_k^i -) \quad \text{for } t \in [\tau_k^i, \tau_{k+1}^i),$$

where u_k^i are independent, uniformly distributed over $\{1, 2, \dots, r\}$.

$i = 1, 2, \dots, r$, $k = 0, 1, 2, \dots$ and $0 = \tau_0^i < \tau_1^i < \tau_2^i < \dots$ are random variables such that $(\tau_{k+1}^i - \tau_k^i)$ are independent, exponentially distributed with respective means $(r-1)/r\alpha\gamma_i$, $i=1, 2, \dots, r$.

Proof: We only prove the case $r = 2$, the general case running along the same lines. By formula (3.1.2), it suffices to show that

$\lim_{n \rightarrow \infty} {}^n(f, \pi)_0 = {}^\infty(f, \pi)_0$ implies $\lim_{n \rightarrow \infty} {}^n(f, \pi)_t = {}^\infty(f, \pi)_t$ almost surely, for each fixed $t \in [0, \infty)$, where for each $n = \infty, 1, 2, 3, \dots$ ${}^n(f, \pi)_t = ({}^n f_t, {}^n \pi_t)$ denotes the dual of $\tilde{Y}_n(t)$; that is, when $n = 1, 2, 3, \dots$, referring to the construction in section (3.4),

$${}^n f_t := \sum_{k=0}^{\infty} I_{[\tau_k^n, \tau_{k+1}^n)}(t) {}^n h_{t-\tau_k^n}^k,$$

where $\tau_k^n := \frac{\varepsilon(1)}{n^2 \varepsilon(n)} \tau_k^1$ for $k = 0, 1, 2, \dots$ and ${}^n h_t^k$ is obtained by

scaling $\gamma_i^n := n^2 \gamma_i^1$ in the construction of R_t . The proof of the

convergence of the finite-dimensional distribution then follows by an inductive argument.

Case a): Clearly, $\|E_{n(f, \pi)_0} \sum_{k=1}^{\infty} I_{[\tau_k^n, \tau_{k+1}^n)}(t) {}^n h_{t-\tau_k^n}^k\| < \dots$

$\|{}^n f_0\| P(\tau_1^n \leq t)$, which goes to 0 as n goes to ∞ . Define

$$\sigma^n := \inf \{t > 0 : {}^n f_t \in C(S^2)\}.$$

$$\lim_{n \rightarrow \infty} E_{n(f, \pi)_0} F(\tilde{Y}_n(0), {}^n(f, \pi)_t) = \lim_{n \rightarrow \infty} E_{n(f, \pi)_0} (F(\tilde{Y}_n(0), {}^n(f, \pi)_t) : \sigma^{\pi_0}_{\tau_1^n})$$

$$= \lim_{n \rightarrow \infty} E_{n(f, \pi)_0} F(\tilde{Y}_n(0), ({}^n h_t^0, {}^n \pi_0))$$

since $\lim_{n \rightarrow \infty} \sigma^n = 0$ and $\lim_{n \rightarrow \infty} \tau_1^n = \infty$. This last term equals

$$= \lim_{n \rightarrow \infty} F(\tilde{Y}_n(0), (E_{n f_0} {}^n h_t^0, {}^n \pi_0)).$$

If we denote by ${}^n R_t$ that operator (involved in the construction of ${}^n h_t^0$) corresponding to that population with diffusion rate n ,

then, since $\lim_{s \rightarrow 0} \|f_s - f_0\| = 0$ implies $\lim_{s \rightarrow 0} \|H_{t+s}^m f_s - H_t^m f_0\| = 0$, we have

$$\lim_{n \rightarrow \infty} \|{}^n R_t {}^n f_0 - H_{2\eta t}^1 \circ \phi_{u_2}^2 \circ \phi_{u_3}^3 \circ \dots \circ \phi_{u_m}^m ({}^\infty f_0)\| = 0 \text{ almost surely, so}$$

$$\lim_{n \rightarrow \infty} \|{}^n h_t^0 - H_{2\eta_1 t}^1 \circ H_{2\eta_2 t}^1 ({}^\infty f_0 \circ D_{\sigma \pi_0})\| = 0 \text{ almost surely,}$$

where $H_{2,1}^1$ acts on the first variable of ${}^\infty f_0 \circ D_{\sigma \pi_0}$, $H_{2,2}^1$ on the second one. But, $\tilde{Y}_n(0) \Rightarrow \tilde{Y}_\infty(0)$ and $\lim_{n \rightarrow \infty} {}^n \pi_0 = {}^\infty \pi_0$, so

$$\lim_{n \rightarrow \infty} E_{n(f, \pi)_0} F(\tilde{Y}_n(0), ({}^n h_t^0, {}^n \pi_0)) = F(\tilde{Y}_\infty(0), (H_{2\eta_1 t}^1 \circ H_{2\eta_2 t}^1 ({}^\infty f_0 \circ D_{\sigma \pi_0}), {}^\infty \pi_0)).$$

By example (3.1.3) with a modified time scale, this is indeed the duality equation for process $\tilde{Y}_\infty(t)$.

Case b): The situation is similar to case a), as $\lim_{n \rightarrow \infty} \sigma^n = 0$, but there is an added difficulty: here, $\tau^n = \inf \{t > 0 : {}^n f_t \in C(S)\}$ is such that $\lim_{n \rightarrow \infty} \tau^n = \tau^\infty$, a random variable, whereas in a), because there $\tau^n \geq \tau_1^n$, $\lim_{n \rightarrow \infty} \tau^n = \infty$. We therefore need to condition on the

value of τ^∞ :

$$\lim_{n \rightarrow \infty} E_{n(f, \pi)_0} (F(\tilde{Y}_n(0), n(f, \pi)_t) \mid \tau^\infty > t) = F(\tilde{Y}_\infty(0), (H_{2\eta_1 t}^1 \circ H_{2\eta_2 t}^1 (f_0 \circ D_{\infty, \pi_0}), \pi_0)),$$

on the set $\{\tau^\infty > t\}$ because then we are still in case a). Otherwise, define $g \in C(S^2)$ by

$$g = H_{2\eta_1 \tau^\infty}^1 \circ H_{2\eta_2 \tau^\infty}^1 (f_0 \circ D_{\infty, \pi_0}) \text{ so that}$$

$$\lim_{n \rightarrow \infty} E_{n(f, \pi)_0} (F(\tilde{Y}_n(0), n(f, \pi)_t) \mid \tau^\infty < t) = H_{2\eta(t-\tau)}^1 \circ g(0, 0)$$

where $\tau = \tau^\infty$ and, $n = n_1$ or n_2 , depending on the variable doing the jump at time τ . Therefore, up to time τ^∞ , the dual is that of the measure-valued process $\tilde{Y}_\infty$ of case a), while after time τ^∞ the dual is just a heat kernel in d -dimension. But, if we write down the generator of process $\tilde{Y}_\infty$ defined in the statement of the theorem, we get:

$$J_\infty F_{(f, \pi)}(\mu) := F(\mu, ([\eta_1 \Delta_1 + \eta_2 \Delta_2](f \circ D_\pi), \pi)) + \\ + \sum_{j=1}^2 \sum_{p=1}^2 \gamma_{\pi(j)} (F(\mu, (f, E_{jp} \pi)) - F(\mu, (f, \pi))).$$

But this is just the generator of the dual described above. $\square$

APPENDIX

We give here a characterization of tightness for sequences in $P(D([0, \infty); R(S)^T))$ where S is a compact metric space. For any metric space (E, d) , recall that $\tilde{\rho}$ denotes the Skorohod metric on $D([0, \infty); E)$ (see section (1:2)).

A.1 Lemma: $\lim_{n \rightarrow \infty} \tilde{\rho}(\omega_n, \omega) = 0$ if and only if, for any sequence $\{t_n\} \subset [0, \infty)$ such that $\lim_{n \rightarrow \infty} t_n = t \in [0, \infty)$, the following three conditions hold:

- (i) $\lim_{n \rightarrow \infty} d(\omega_n(t_n), \omega(t)) \wedge d(\omega_n(t_n), \omega(t-)) = 0;$
- (ii) If $\lim_{n \rightarrow \infty} d(\omega_n(t_n), \omega(t)) = 0$ and $s_n \leq t_n$ for all $n \in \mathbb{Z}^+$ with $\lim_{n \rightarrow \infty} s_n = t$, then $\lim_{n \rightarrow \infty} d(\omega_n(s_n), \omega(t)) = 0;$
- (iii) If $\lim_{n \rightarrow \infty} d(\omega_n(t_n), \omega(t-)) = 0$ and $0 \leq s_n \leq t_n$ for all $n \in \mathbb{Z}^+$ with $\lim_{n \rightarrow \infty} s_n = t$, then $\lim_{n \rightarrow \infty} d(\omega_n(s_n), \omega(t-)) = 0.$

Proof: See Ethier and Kurtz [EK], pp. 125-127. { }

As a consequence of this result, if $f: E_1 \rightarrow E_2$ is a continuous function from one metric space E_1 into another E_2 , then $\tilde{f}: D([0, \infty); E_1) \rightarrow D([0, \infty); E_2)$, defined by $\tilde{f}(\omega)(t) := f(\omega(t))$ at each $t \in [0, \infty)$, is also continuous. Another consequence is the following:

A.2 Proposition: A subset $A \subset D([0, \infty); \mathbb{R}^d)$ is relatively compact

if and only if the subsets $\phi_{ij}A$ and $\phi_{ij}A$ are relatively compact in $D([0, \infty); \mathbb{R})$, for all $i, j = 1, 2, \dots, d$, where $\phi_i(x) := (x(\cdot), e_i)$, $\phi_{ij}(x) := \phi_i(x) + \phi_j(x)$ and $e_1, e_2, \dots, e_d$ is an orthonormal basis in $\mathbb{R}^d$.

Proof: The condition is necessary since ϕ_1, ϕ_{ij} , $i, j = 1, \dots, d$ are continuous. For the sufficiency, any sequence $\{x_n\} \subset A$ contains a subsequence (labelled $\{x_n\}$ as well here) such that, for some $x \in D([0, \infty); \mathbb{R}^d)$, $\lim_{n \rightarrow \infty} \tilde{\rho}(\phi_1(x_n), \phi_1(x)) = 0$ and

$\lim_{n \rightarrow \infty} \rho(\phi_{ij}(x_n), \phi_{ij}(x)) = 0$ for all $i, j = 1, 2, \dots, d$. To prove that

$\lim_{n \rightarrow \infty} \tilde{\rho}(x_n, x) = 0$, we use lemma (A.1). Let $a_1^n := |\phi_1(x_n(t_n)) - \phi_1(x(t))|$,

$b_1^n := |\phi_1(x_n(t_n)) - \phi_1(x(t-))|$, defining a_{ij}^n, b_{ij}^n

the same way with ϕ_{ij} in place of ϕ_1 . We must prove that

$\lim_{n \rightarrow \infty} a_i^n \wedge b_i^n = 0$, and $\lim_{n \rightarrow \infty} a_{ij}^n \wedge b_{ij}^n = 0$ imply $\lim_{n \rightarrow \infty} \left(\sum_{i=1}^d a_i^n \right) \wedge \left(\sum_{i=1}^d b_i^n \right) = 0$ in order to prove condition (i), the other two

conditions following easily. Now,

conditions following easily. Now,

$$\left(\sum_{i=1}^d a_i^n \right) \wedge \left(\sum_{i=1}^d b_i^n \right) \leq \sum_{i,j=1}^d (a_i^n \wedge b_j^n) \leq \sum_{i,j=1}^d (a_{i+j}^n \wedge b_{i+j}^n)$$

so it suffices to prove that $\lim_{n \rightarrow \infty} (a_{i+j}^n \wedge b_{i+j}^n) = 0$. This is

done case by case, by splitting the sequence $\{x_n\}$ into eight

subsequences, each of which satisfies $\lim_{n \rightarrow \infty} a_i^n = 0$ (or $\lim_{n \rightarrow \infty} b_i^n = 0$)

and $\lim_{n \rightarrow \infty} a_j^n = 0$ (or $\lim_{n \rightarrow \infty} b_j^n = 0$) and $\lim_{n \rightarrow \infty} a_{ij}^n = 0$ (or

$\lim_{n \rightarrow \infty} b_{ij}^n = 0$). For instance, if

$$\lim_{n \rightarrow \infty} |\phi_1(x_n(t_n)) - \phi_1(x(t))| = 0,$$

$$\lim_{n \rightarrow \infty} |\phi_j(x_n(t_n)) - \phi_j(x(t-))| = 0,$$

$$\lim_{n \rightarrow \infty} |(\phi_1(x_n(t_n)) - \phi_1(x(t))) + (\phi_j(x_n(t_n)) - \phi_j(x(t)))| = 0,$$

then clearly $\phi_j(x(t-)) = \phi_j(x(t))$ so because of the second equation we get $\lim_{n \rightarrow \infty} (a_1^n + a_j^n) = 0$. The other cases run likewise.

Note that the relative compactness of $\{A_i\}$ for all $i=1,2,\dots,d$ alone is not sufficient to ensure the relative compactness of A :

let $(x(t), e_1) := \delta_{[1,\infty)}(t)$ for all $i=1,2,\dots,d$ and

$$(x_n(t), e_1) = \begin{cases} \delta_{[1,\infty)}(t) & \text{if } i=1 \\ \delta_{[1+1/n,\infty)}(t) & \text{if } i \neq 1. \end{cases} \quad \text{For } t_n - t = 1 \text{ for all } n$$

$n \in \mathbb{Z}^+$, condition (1) holds for each coordinate, but not for x_n and x .

A.3 Theorem: Let E be a compact metric space and H a dense subset of $C(E)$. A sequence $\{P_n : n \in \mathbb{Z}^+\} \subset P(D([0,\infty); E))$ is tight if and only if, for each $f \in H$, the sequence $\{P_n \circ f^{-1} : n \in \mathbb{Z}^+\} \subset P(D([0,\infty); \mathbb{R}))$ is tight.

Proof: See Ethier and Kurtz [EK], p. 142. $\square$

The next theorem states that, when $E = P(S)^r$ where S is a compact metric space, we can select a much smaller collection of

functions H in theorem (A.3). Let $M \subset C(S)$ denote some countable convergence-determining set for $P(S)$ and write $M = \{f_k : k \in \mathbb{Z}^+\}$. For any $f \in C(S)$, define $F_f^1 : P(S)^r \rightarrow \mathbb{R}$ by $F_f^1(\mu) := \int_S f d\mu_1$ for each $i = 1, 2, \dots, r$.

A.4 Theorem: A sequence $\{P_n : n \in \mathbb{Z}^+\} \subset P(D([0, \infty) : P(S)^r))$ is tight if and only if, for every choice of $i, j = 1, 2, \dots, r$ and every choice of $f, g \in M$, the sequence $\{P_n \circ (F_f^1 + F_g^j)^{-1} : n \in \mathbb{Z}^+\} \subset P(D([0, \infty) : \mathbb{R}))$ is tight.

Proof: Necessity is clear. Sufficiency will be proved, using theorem (1.2.4), if we can exhibit, for each $\epsilon > 0$, a set $A_\epsilon \subset D([0, \infty) : P(S)^r)$ such that

$$A.5 \quad P_n(A_\epsilon^c) \leq \epsilon \quad \text{for all } n \in \mathbb{Z}^+$$

and

$$A.6 \quad \limsup_{\delta \rightarrow 0} \sup_{\mu \in A_\epsilon} W'(\mu, \delta, T) \leq \eta \quad \text{for all } \eta > 0 \text{ and } T > 0,$$

where

$$W'(\mu, \delta, T) := \inf_{\{t_m\}} \max_m \sup_{s, t \in [t_{m-1}, t_m]} \sum_{i=1}^r \sum_{k=1}^{\infty} 2^{-k} (1 \wedge |\int f_k d\mu_i(s) - \int f_k d\mu_i(t)|),$$

the infimum taken over all finite collections $0 = t_0 < t_1 < \dots < t_{m-1} < T \leq t_m < \infty$ with $(t_m - t_{m-1}) \geq \delta$ for all $m = 1, 2, 3, \dots, n$.

For convenience, denote $\tilde{F}_{f_k}^i$ by ψ_k^i and $\tilde{F}_{f_k}^i + \tilde{F}_{f_l}^j$ by $\psi_{k,l}^{i,j}$.

By hypothesis, there exists compact sets K_k^i and $K_{k,l}^{i,j}$ in

$D([0, \infty): \mathbb{R})$ such that, for all $n \in \mathbb{Z}^+$,

$$P_n \circ (\psi_k^i)^{-1} (K_k^i)^c \leq \varepsilon 2^{-i-k-1} \quad \text{and} \quad P_n \circ (\psi_{k,\ell}^{1,j})^{-1} (K_{k,\ell}^{1,j})^c \leq \varepsilon 2^{-1-j-k-\ell-1}.$$

Thus (A.5) is satisfied with $A_{i,k} := (\cap_{i,k} (\psi_k^i)^{-1} K_k^i) \cap$

$(\cap_{i,j,k,\ell} (\psi_{k,\ell}^{1,j})^{-1} K_{k,\ell}^{1,j})$. We must verify (A.6). Choose p such that

$$\sum_{j=p+1}^{\infty} 2^{-j} < \eta/2r \quad \text{and define}$$

$$A_\varepsilon^p := y(A_\varepsilon) \quad \text{where} \quad y(\omega) := (\psi_k^1(\omega))_{i=1,2,\dots,r; k=1,2,\dots,p} \in D([0, \infty): \mathbb{R}^{pr}).$$

Now, using proposition (A.2), for every $\alpha, \beta = 1, 2, \dots, pr$, we have

$$\Phi_{\alpha\beta}^p \subset K_k^1, \quad \text{for some choice of } i=1, 2, \dots, r \text{ and } k=1, 2, \dots, p; \text{ and}$$

$$\Phi_{\alpha\beta}^p \subset K_{k,\ell}^{1,j}, \quad \text{for some choice of } i,j=1, 2, \dots, r \text{ and } k,\ell=1, 2, \dots, p;$$

so $A_\varepsilon^p \subset D([0, \infty): \mathbb{R}^{pr})$ is compact. By theorem (1.2.4),

$$\lim_{\delta \rightarrow 0} \sup_{y \in A_\varepsilon^p} W(y, \delta, T) = 0 \quad \text{so choose } \delta^* \text{ so that, for each } y \in A_\varepsilon^p,$$

there exists a partition $\{t_m^y\}$ with $(t_m^y - t_{m-1}^y) \leq \delta^*$ such that

$$G(y) := \max_m \sup_{s, t \in [t_{m-1}^y, t_m^y]} |y(s) - y(t)| < \eta/2r.$$

Then, for all $\mu \in A_\varepsilon$,

$$\begin{aligned} W(\mu, \delta^*, T) &= \inf_{\{t_m\}} \max_m \sup_{s, t \in [t_{m-1}, t_m]} \sum_{i=1}^r \sum_{k=1}^p 2^{-k} (1 \wedge |\psi_k^i(\mu(s)) - \psi_k^i(\mu(t))|) \\ &\leq \eta/2 + \sum_{i=1}^r \sum_{k=1}^p 2^{-k} (1 \wedge G(y(\mu))) < \eta. \quad \square \end{aligned}$$

BIBLIOGRAPHY

- [AL] Aldous, D., Exchangeability and Related Topics, in Lecture Notes in Mathematics 1117, pp. 1-199, Springer-Verlag 1985.
- [BI] Billingsley, P., Probability and Measure, Wiley 1979.
- [DA1] Dawson, D.A., Critical Dynamics and Fluctuations for a Mean-Field Model of Cooperative Behavior, Journal of Statistical Physics, Vol. 31, No. 1, 1983.
- [DA2] Dawson, D.A., Measure-Valued Processes: Construction, Qualitative Behavior and Stochastic Geometry, in Lecture Notes in Mathematics 1212, p. 69-93, Springer-Verlag 1986.
- [DH1] Dawson, D.A. and Hochberg, K.J., Wandering Random Measures in the Fleming-Viot Model, Annals of Probability, Vol. 10, No. 3, pp. 554-580, 1982.
- [DH2] Dawson, D.A. and Hochberg, K.J., Function-Valued Duals for Measure-Valued Processes and Applications, A.M.S. Contemporary Mathematics, Vol. 41, pp. 55-69, 1985.
- [DK] Dawson, D.A. and Kurtz, T.G., Applications of Duality to Measure-Valued Processes, in Lecture Notes in Control and Information Sciences 42, pp. 177-191, Springer-Verlag 1982.
- [EK] Ethier, S.N. and Kurtz, T.G., Markov Processes: Characterization and Convergence, Wiley 1986.
- [FV] Fleming, W.H. and Viot, M., Some Measure-Valued Markov Processes in Population Genetics Theory, Indiana University Mathematics Journal, Vol. 28, pp. 817-843, 1979.

- [JM] Joffe, A. et Métivier, M., Weak Convergence of Sequences of Semimartingales with Applications to Multitype Branching Processes, Advances in Applied Probability, Vol. 18, pp. 20-65, 1986.
- [LA] Lang, S., Real Analysis, Addison-Wesley 1969.
- [LE] Léonard, C., Une Loi des Grands Nombres pour des Systèmes de Diffusions avec Interaction et Coefficients Non Bornés, Annales de l'Institut Henri Poincaré, Vol. 22, No. 2, pp. 237-262, 1986.
- [MP] Métivier, M. and Pellaumail, J., Stochastic Integration, Academic Press 1980.
- [OE] Oelschläger, K., A Martingale Approach to the Law of Large Numbers for Weakly Interacting Stochastic Processes, Annals of Probability, Vol. 12, No. 2, pp. 458-479, 1984.
- [OK] Ohta, T. and Kimura, M., A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population, Genetical Research, Vol. 22, pp. 201-204, 1973.
- [PA] Parthasarathy, K.R., Probability Measures in Metric Spaces, Academic Press 1967.
- [SV] Stroock, D.W. and Varadhan, S.R.S., Multidimensional Diffusion Processes, Springer-Verlag 1979.

END

20/06/88

FIN