Part II

Edwin Perkins: Dawson–Watanabe Superprocesses and Measure-valued Diffusions

E. Bolthausen, E. Perkins, A. van der Vaart: LNM 1781, P. Bernard (Ed.), pp. 125–329, 2002.

Table of Contents

Outline of Lectures at St. Flour

1. Introduction
2. Particle Sytems and Tightness (II.1-II.4)
3. The Martingale Problem and Non-linear Equation (II.4-II.8)
4. Path Properties of the Support of Super-Brownian Motion (III.1-III.3)
5. Polar Sets (III.5-III.6)
6. Interactive Drifts (IV)
7. Spatial Interactions 1.
   Stochastic Integration on Trees and a Strong Equation (V.1-V.3)
8. Spatial Interactions 2.
   Pathwise Existence & Uniqueness, and the Historical Martingale Problem (V.4-V.5)
9. Interacting Particle Systems 1. The Voter Model
10. Interacting Particle Systems 2. The Contact Process

Note. A working document with Ted Cox and Rick Durrett was distributed to provide background material for lectures 9 and 10.

Glossary of Notation

$\alpha \sim t$	$\lvert\alpha\rvert/N \le t < (\lvert\alpha\rvert+1)/N = \zeta^\alpha$, i.e., α labels a particle alive at time t
$\lvert A\rvert$	Lebesgue measure of A
A^δ	the set of points less than a distance δ from the set A
$A^g\phi$	$A\phi + g\phi$
$\hat{A}$	weak generator of path-valued Brownian motion–see Lemma V.2.1
$\bar{A}_{\tau,m}$	see Proposition V.2.6
$\vec{A}$	generator of space-time process–see prior to Proposition II.5.8
$\xrightarrow{bp}$	bounded pointwise convergence
$b\mathcal{E}$	the set of bounded $\mathcal{E}$-measurable functions
BSMP	a cadlag strong Markov process with $x \mapsto P^x(A)$ Borel measurable for each measurable A in path space
$\mathcal{B}(E)$	the Borel σ-field on E
C	$C(\mathbb{R}^d)$
$C(E)$	continuous E-valued functions on $\mathbb{R}_+$ with the topology of uniform convergence on compacts
$C_b(E)$	bounded continuous E-valued functions on $\mathbb{R}_+$ with the supnorm topology
$C_b^k(\mathbb{R}^d)$	functions in $C_b(\mathbb{R}^d)$ with bounded continuous partials of order k or less
$C_b^\infty(\mathbb{R}^d)$	functions in $C_b(\mathbb{R}^d)$ with bounded continuous partials of any order
$C_K(E)$	continuous functions with compact support on E with the supnorm topology
$C_\ell(E)$	continuous functions on a locally compact E with a finite limit at ∞
$C(g)(A)$	the g-capacity of a set A–see prior to Theorem III.5.2
$\mathcal{C}$	Borel σ-field for $C = C(\mathbb{R}^d)$
$\mathcal{C}_t$	sub-σ-field of $\mathcal{C}$ generated by coordinate maps y_s, $s \le t$
$\stackrel{\mathcal{D}}{\equiv}$	equal in distribution
$D(E)$	the space of cadlag paths from $\mathbb{R}_+$ to E with the Skorokhod J_1 topology
D^s	the set of paths in $D(E)$ which are constant after time s
D_{fd}	smooth functions of finitely many coordinates on $\mathbb{R}_+ \times C$ –see Example V.2.8
$D(n,d)$	space of $\mathbb{R}^{n\times d}$-valued integrands–see after Proposition V.3.1
Δ	cemetary state added to E as a discrete point
$\mathcal{D}(A)$	domain of the weak generator A–see II.2 and Proposition II.2.2
$\mathcal{D}(\vec{A})_T$	domain of weak space-time generator–see prior to Proposition II.5.7
$\mathcal{D}(\hat{A})$	domain of the weak generator for path-valued Brownian motion –see Lemma V.2.1
$\mathcal{D}(\Delta/2)$	domain of the weak generator of Brownian motion–see Example II.2.4
$\mathcal{D}$	the Borel σ-field on the Skorokhod space $D(E)$
$(\mathcal{D}_t)_{t\ge 0}$	the canonical right-continuous filtration on $D(E)$
$e_\phi(W)$	$e^{-W(\phi)}$
$\mathcal{E}$	the Borel σ-field on E
$\mathcal{E}_+$	the non-negative $\mathcal{E}$-measurable functions
$\hat{E}$	$\{(t, y(\cdot \wedge t)) : y \in D(E), t \ge 0\}$
$f_\beta(r)$	r^β if $\beta > 0$, $(\log 1/r)^{-1}$ if $\beta = 0$
$\hat{\mathcal{F}}$	$\mathcal{F} \times \mathcal{B}(C(\mathbb{R}^d))$
$\hat{\mathcal{F}}_t$	$\mathcal{F}_t \times \mathcal{C}_t$
$\hat{\mathcal{F}}_t^*$	the universal completion of $\hat{\mathcal{F}}_t$

$\mathcal{F}_X$	the Borel σ-field on Ω_X
$g_\beta(r)$	$r^{-\beta}$ if $\beta > 0$, $1 + (\log 1/r)^+$, if $\beta = 0$ and 1, if $\beta < 0$
$G_\varepsilon\phi$	see (IV.3.4)
$G(f,t)$	$\int_0^t \sup_x P_s f(x) ds$
$G(X)$	$\cup_{\delta>0}\mathrm{cl}\{(t,x) : t \geq \delta,\ x \in S(X_t)\}$, the graph of X
$h-m$	the Hausdorff h-measure–see Section III.3
$h(r)$	Lévy's modulus function $(r\log(1/r))^{1/2}$
$h_d(r)$	$r^2 \log^+ \log^+ 1/r$ if $d \geq 3$, $r^2(\log^+ 1/r)(\log^+ \log^+ \log^+ 1/r)$ if $d = 2$
$H_t^{s,y}$	the H_t measure of $\{w : w = y \text{ on } [0,s]\}$, $s \leq t$, $y(\cdot) = y(\cdot \wedge s)$
$\overline{\mathcal{H}}^{bp}$	the bounded pointwise closure of $\mathcal{H}$
$\mathcal{H}_+$	the set of non-negative functions in $\mathcal{H}$
I	$\bigcup_{n=0}^{\infty} \mathbb{N}^{\{0,\dots,n\}} = \{(\alpha_0,\dots,\alpha_n) : \alpha_i \in \mathbb{N}, n \in \mathbb{Z}_+\}$
IBSMP	time inhomogeneous Borel strong Markov process–see after Lemma II.8.1
$I(f,t)$	stochastic integral of f on a Brownian tree–see Proposition V.3.2
$\mathcal{K}$	the compact subsets of $\mathbb{R}^d$
Lip_1	Lipschitz continuous functions with Lipschitz constant and supnorm ≤ 1
(LE)	Laplace functional equation–see prior to Theorem II.5.11
$(LMP)_\nu$	local martingale problem for Dawson-Watanabe superprocess with initial law ν–see prior to Theorem II.5.1
$L_t(X)$	the collision local time of $X = (X^1, X^2)$–see prior to Remarks IV.3.1
$\log^+(x)$	$(\log x) \vee e^e$
$L_W(\phi)$	the Laplace functional of the random measure W, i.e., $E(e^{-W(\phi)})$
$\mathcal{L}^2_{\mathrm{loc}}$	see after Lemma II.5.2
$M_1(E)$	space of probabilities on E with the topology of weak convergence
$M_F(E)$	the space of finite measures on E with the topology of weak convergence
$M_F^t(D)$	the set of finite measures on $D(E)$ supported by paths which are constant after time t
$\mathcal{M}_F$	the Borel σ-field on $M_F(E)$
$\mathcal{M}_{\mathrm{loc}}$	the space of continuous $(\mathcal{F}_t)$-local martingales starting at 0
(ME)	mild form of the nonlinear equation–see prior to Theorem II.5.11
$(MP)_{X_0}$	martingale problem for Dawson-Watanabe superprocess with initial state X_0 –see Proposition II.4.2
$\hat{\Omega}$	$\Omega \times C(\mathbb{R}^d)$
$\Omega_H[\tau,\infty)$	$\Big\{H_\cdot \in C\Big([\tau,\infty), M_F\big(D(E)\big)\Big) : H_t \in M_F^t(D) \quad \forall t \geq \tau\Big\}$
Ω_H	$\Omega_H[0,\infty)$
Ω_X	the space of continuous $M_F(E)$-valued paths
Ω_D	the space of cadlag $M_F(E)$-valued paths
$p_t(x)$	standard Brownian density
$p_t^x(y)$	$p_t(x-y)$
$\mathcal{P}$	the σ-field of $(\mathcal{F}_t)$-predictable subsets of $\mathbb{R}_+ \times \Omega$
$P_t^g\phi(x)$	$E^x\Big(\phi(Y_t)\exp\Big\{\int_0^t g(Y_s)\,ds\Big\}\Big)$
$\mathbb{P}_{X_0}$	the law of the DW superprocess on $(\Omega_X, \mathcal{F}_X)$ with initial state X_0 –see Theorem II.5.1
$\mathbb{P}_\nu$	the law of the DW superprocess with initial law ν

$\hat{\mathbb{P}}_T$	the normalized Campbell measure associated with K_T, i.e., $\hat{\mathbb{P}}_T(A\times B)=\mathbb{P}(1_A K_T(B))/m(1)$
(PC)	$x\mapsto P^x$ is continuous
(QLC)	quasi-left continuity, i.e., Y is a Hunt process–see Section II.2
$\mathbb{Q}_{\tau,m}$	the law of the historical process starting at time τ in state m–see Section II.8
$\mathcal{R}(I)$	$\bigcup_{t\in I} S(X_t)$, the range of X on I
$\overline{\mathcal{R}}(I)$	$\overline{\mathcal{R}(I)}$ is the closed range of X on I
$\mathcal{R}$	$\bigcup_{\delta>0}\overline{\mathcal{R}}\big([\delta,\infty)\big)$ is the range of X.
$S(\mu)$	the closed support of a measure μ
S_t	$S(X_t)$
$\mathcal{S}$	simple $\mathcal{P}\times\mathcal{E}$-measurable integrands–see after Lemma II.5.2
(SE)	strong form of nonlinear equation–see prior to Theorem II.5.11
$\underline{t}$	$[Nt]/N$
$\mathcal{T}_b$	bounded $(\mathcal{F}_t)_{t\ge\tau}$-stopping times
$\stackrel{ucb}{\to}$	convergence on E which is uniform on compacts and bounded on E
U_λ	the λ resolvent of a Markov process
$\stackrel{w}{\Rightarrow}$	weak convergence of finite (usually probability) measures
W_t	the coordinate maps on $D(\hat{E})$
$y/s/w$	the path equaling y up to s and $w(t-s)$ thereafter
$y^t(\cdot)$	$y(t\wedge\cdot)$
ζ_α	the lifetime of the α^{th} branch–see after Remark II.3.2

I. Introduction

Over the years I have heard a number of complaints about the impenetrable literature on measure-valued branching processes or Dawson-Watanabe superprocesses. These concerns have in part been addressed by some recent publications including Don Dawson's St. Flour notes (Dawson (1993)), Eugene Dynkin's monograph (Dynkin (1994)) and Jean-Francois Le Gall's ETH Lecture Notes (Le Gall (1999)). Nonetheless, one still hears that several topics are only accessible to experts. However, each time I asked a colleague what topics they would like to see treated in these notes, I got a different suggestion. Although there are some other less flattering explanations, I would like to think the lack of a clear consensus is a reflection of the large number of different entry points to the subject. The Fleming-Viot processes, used to model genotype frequencies in population genetics, arise by conditioning the total mass of a superprocess to be one (Etheridge and March (1991)). When densities exist (as for super-Brownian motion in one spatial dimension) they typically are solutions of parabolic stochastic pde's driven by a white noise and methods developed for their study often have application to large classes of stochastic pde's (e.g. Mueller and Perkins (1992), Krylov (1997b), Mytnik (1998) and Section III.4). Dawson-Watanabe superprocesses arise as scaling limits of interacting particle systems (Cox, Durrett and Perkins (1999, 2000)) and of oriented percolation at criticality (recent work of van der Hofstad and Slade (2000)). Rescaled lattice trees above eight dimensions converge to the integral of the super-Brownian cluster conditioned to have mass one (Derbez and Slade (1998)). There are close connections with class of nonlinear pde's and the interaction between these fields has led to results for both (Dynkin and Kuznetsov (1996,1998), Le Gall (1999) and Section III.5). They provide a rich source of exotic path properties and an interesting collection of random fractals which are amenable to detailed study (Perkins (1988), Perkins and Taylor (1998), and Chapter III).

Those looking for an overview of all of these developments will not find them here. If you are looking for "the big picture" you should consult Dawson (1993) or Etheridge (2000). My goal in these notes is two-fold. The first is to give a largely self-contained graduate level course on what John Walsh would call "the worm's-eye view of superprocesses". The second is to present some of the topics and methods used in the study of interactive measure-valued models.

Chapters II and III grew out of a set of notes I used in a one-semester graduate course on Superprocesses. A version of these notes, recorded by John Walsh in a legible and accurate hand, has found its way to parts of the community and in fact been referenced in a number of papers. Although I have updated parts of these notes I have not tried to introduce a good deal of the more modern machinery, notably Le Gall's snake and Donnelly and Kurtz's particle representation. In part this is pedagogical. I felt a direct manipulation of branching particle systems (as in II.3,II.4) allows one to quickly gain a good intuition for superprocesses, historical processes, their martingale problems and canonical measures. All of these topics are described in Chapter II. In the case of Le Gall's snake, Le Gall (1999) gives an excellent and authoritative treatment. Chapter III takes a look at the qualitative properties of Dawson-Watanabe superprocesses. Aside from answering a number of natural questions, this allows us to demonstrate the effectiveness of the various tools used to study branching diffusions including the related nonlinear parabolic pde, historical processes, cluster representations and the martingale problem. Although many of the results presented here are definitive, a number of open problems and conjectures are stated. Most of the Exercises in these Chapters play a crucial role in the presentation and are highly recommended.

My objective in Chapters II and III is to present the basic theory in a middling degree of generality. The researcher looking for a good reference may be disappointed that we are only considering finite variance branching mechanisms, finite initial conditions and Markov spatial motions with a semigroup acting on the space of bounded continuous functions on a Polish space E. The graduate student learning the subject or the instructor teaching a course, may be thankful for the same restrictions. I have included such appendages as location dependent branching and drifts as they motivate some of the interactions studied in Chapters IV and V. Aside from the survey in Section III.7, every effort has been made to provide complete proofs in Chapters II and III. The reader is assumed to have a good understanding of continuous parameter Markov processes and stochastic calculus–for example, the first five Chapters of Ethier and Kurtz (1986) provide ample background. Some of the general tightness results for stochastic processes are stated with references (notably Lemma II.4.5 and (II.4.10), (II.4.11)) but these are topics best dealt with in another course. Finally, although the Hausdorff measure and polar set results in Sections III.3 and III.5 are first stated in their most general forms, complete proofs are then given for slightly weaker versions. This means that at times when these results are used, the proofs may not be self-contained in the critical dimensions (e.g. in Theorem III.6.3 when $d = 4$).

A topic which was included in the original notes but not here is the Fleming-Viot process (but see Exercise IV.1.2). The interplay between these two runs throughout Don Dawson's St. Flour notes. The reader should really consult the article by Ethier and Kurtz (1993) to complete the course.

The fact that we are able to give such a complete theory and description of Dawson-Watanabe superprocesses stems from the strong independence assumptions underlying the model which in turn produces a rather large tool kit for their study. Chapters IV and V study measure-valued processes which may have state-dependent drifts, spatial motions and branching rates (the latter is discussed only briefly). All of the techniques used to study ordinary superprocesses become invalid or must be substantially altered if such interactions are introduced into the model. This is an ongoing area of active research and the emphasis here is on introducing some approaches which are currently being used. In Chapter IV, a competing species model is used to motivate careful presentations of Dawson's Girsanov theorem for interactive drifts and of the construction of collision local time for a class of measure-valued processes. In Chapter V, a strong equation driven by a historical Brownian motion is used to model state dependent spatial motions. Section IV.4 gives a discussion of the competing species models in higher dimensions and Section V.5 describes what is known about the martingale problems for these spatial interactions. The other sections in these chapters are again self-contained with complete arguments.

There are no new results contained in these notes. Some of the results although stated and well-known are perhaps not actually proved in the literature (e.g. the disconnectedness results in III.6) and some of the proofs presented here are, I hope, cleaner and shorter. I noticed that some of the theorems were originally derived using nonstandard analysis and I have standardized the arguments (often using the historical process) to make them more accessible. This saddens me a bit as I feel the nonstandard view, clumsy as it is at times, is pedagogically superior and allows one to come up with novel insights.

As one can see from the outline of the actual lectures, at St. Flour some time was spent on rescaled limits of the voter model and the contact process, but these topics have not made it into these notes. A copy of some notes prepared with Ted Cox and Rick Durrett on this subject was distributed at St. Flour and is available from me (or them) upon request. We were trying to unify and extend these results. As new applications are still emerging, I decided it would be better to wait until they find a more definitive form than rush and include them here. Those who have seen earlier versions of these notes will know that I also had planned to include a detailed presentation of the particle representations of Donnelly and Kurtz (1999). In this case I have no real excuse for not including them aside from running out of time and a desire to keep the total number of pages under control. They certainly are one of the most important techniques available for treating interactive measure-valued models and hence should have been included in the second part of these notes.

There a number of people to thank. First the organizers and audience of the 1999 St. Flour Summer School in Probability for an enjoyable and at times exhausting $2\frac{1}{2}$ weeks. A number of suggestions and corrections from the participants has improved these notes. The Fields Institute invited me to present a shortened and dry run of these lectures in February and March, and the audience tolerated some experiments which were not entirely successful. Thanks especially to Siva Athreya, Eric Derbez, Min Kang, George Skoulakis, Dean Slonowsky, Vrontos Spyros, Hanno Treial and Xiaowen Zhou. Most of my own contributions to the subject have been joint and a sincere thanks goes to my co-authors who have contributed to the results presented at St. Flour and who have made the subject so enjoyable for me: Martin Barlow, Ted Cox, Don Dawson, Rick Durrett, Steve Evans, Jean-Francois Le Gall and Carl Mueller. Finally a special thanks to Don Dawson and John Walsh who introduced me to the subject and have provided ideas which can be seen throughout these notes.

II. Branching Particle Systems and Dawson-Watanabe Superprocesses

1. Feller's Theorem

Let $\{X_i^k : i \in \mathbb{N}, k \in \mathbb{Z}_+\}$ be i.i.d. $\mathbb{Z}_+$-valued random variables with mean 1 and variance $\gamma > 0$. We think of X_i^k as the number of offspring of the i^{th} individual in the k^{th} generation, so that $Z_{k+1} = \sum_{i=1}^{Z_k} X_i^k$ (set $Z_0 \equiv 1$) is the size of the $k+1^{\text{st}}$ generation of a Galton-Watson branching process with offspring distribution $\mathcal{L}(X_i^k)$, the law of X_i^k.

We write $a_n \sim b_n$ iff $\lim_{n\to\infty} a_n/b_n = 1$ and let $\stackrel{w}{\Rightarrow}$ denote weak convergence of finite (usually probability) measures.

Theorem II.1.1. (a) (Kolmogorov (1938)) $P(Z_n > 0) \sim 2/n\gamma$ as $n \to \infty$.
(b) (Yaglom (1947)) $P\left(\frac{Z_n}{n} \in \cdot \mid Z_n > 0\right) \stackrel{w}{\Rightarrow} Z$, where Z is exponential with mean $\gamma/2$.

Proof. (a) This is a calculus exercise in generating functions but it will be used on several occasions and so we provide a proof. Let $f_n(t) = E(t^{Z_n})$ for $t \in [0,1]$, and $f(t) = f_1(t)$ (here $0^0 = 1$ as usual). A simple induction shows that f_n is the n-fold composition of f with itself. Then Dominated Convergence shows that f' and f'' are continuous on $[0,1]$, where the appropriate one-sided derivatives are taken at the endpoints. Moreover $f'(1) = E(X_i^k) = 1$ and $f''(1) = \text{var}(X_i^k) = \gamma$. As f is increasing and strictly positive at 0 (the latter because X_i^k has mean 1 and is not constant), we must have

$$0 < f(0) \le f_n(0) \uparrow L \le 1 \text{ and } f(L) = L.$$

Note that $f'(t) = E(Z_1 t^{Z_1 - 1}) < 1$ for $t < 1$ and so the Mean Value Theorem implies that $f(1) - f(t) < 1 - t$ and therefore, $f(t) > t$, for $t < 1$. This proves that $L = 1$ (as you probably already know from the a.s. extinction of the critical branching process).

Set $x_n = f_n(0)$ and

$$(II.1.1) \qquad y_n = n(1 - x_n) = nP(Z_n > 0).$$

A second order Taylor expansion shows that

$$1 - x_{n+1} = f(1) - f(x_n) = 1 - x_n - \frac{f''(z_n)}{2}(1 - x_n)^2 \text{ for some } z_n \in [x_n, 1].$$

Therefore

$$\begin{aligned} y_{n+1} &= (n+1)(1 - x_{n+1}) \\ &= (n+1)[1 - x_n - \frac{f''(z_n)}{2}(1 - x_n)^2] \\ &= (n+1)\left[\frac{y_n}{n} - \frac{y_n^2}{n^2}\frac{f''(z_n)}{2}\right] \\ &= y_n\left[1 + \frac{1}{n}\left(1 - y_n(1 + n^{-1})\frac{f''(z_n)}{2}\right)\right]. \end{aligned}$$

Now let $\gamma_1 < \gamma < \gamma_2$ and $\delta > 0$. Note that $\lim_{n\to\infty} x_n = 1$ and $\lim_{n\to\infty} f''(z_n) = \gamma$, and so we may choose n_0 so that for $n \geq n_0$,

$$(1 - x_n)\gamma_2/2 < \delta, \tag{II.1.2}$$

$$y_n\Big[1 + \frac{1}{n}(1 - y_n\gamma_2/2)\Big] \leq y_{n+1} \leq y_n\Big[1 + \frac{1}{n}(1 - y_n\gamma_1/2)\Big], \tag{II.1.3}$$

and therefore,

$$y_{n+1} > y_n \text{ if } y_n < 2/\gamma_2, \text{ and } y_{n+1} < y_n \text{ if } y_n > 2/\gamma_1. \tag{II.1.4}$$

Claim $y_n > \frac{2}{\gamma_2}(1-\delta)$ eventually. Note first that if $n_1 \geq n_0$ satisfies $y_{n_1} \leq \frac{2}{\gamma_2}(1-\delta/2)$ (there is nothing to prove if no such n_1 exists), then the lower bound in (II.1.3) shows that $y_{n_1+1} \geq y_{n_1}\Big(1 + \frac{\delta}{2n_1}\Big)$. Iterating this bound, we see that there is an $n_2 > n_1$ for which $y_{n_2} > \frac{2}{\gamma_2}(1-\delta/2)$. Now let n_3 be the first $n > n_2$ for which $y_n \leq \frac{2}{\gamma_2}(1-\delta)$ (again we are done if no such n exists). Then $y_{n_3-1} > \frac{2}{\gamma_2}(1-\delta) \geq y_{n_3}$ and so (II.1.4) implies that $y_{n_3-1} \geq \frac{2}{\gamma_2}$. Therefore (II.1.3) shows that

$$\begin{aligned} y_{n_3} &\geq y_{n_3-1}\Big[1 + \frac{1}{n_3-1}(1 - y_{n_3-1}\gamma_2/2)\Big] \\ &\geq \frac{2}{\gamma_2}\Big[1 - \frac{y_{n_3-1}}{n_3-1}\frac{\gamma_2}{2}\Big] \\ &= \frac{2}{\gamma_2}\Big[1 - \frac{(1-x_{n_3-1})\gamma_2}{2}\Big] \\ &> \frac{2}{\gamma_2}(1-\delta), \end{aligned}$$

the last by (II.1.2). This contradicts the choice of n_3 and hence proves the required inequality for $n \geq n_2$. A similar argument shows that $y_n \leq \frac{2}{\gamma_1}(1-\delta)$ eventually. We thus have shown that $\lim_{n\to\infty} y_n = 2/\gamma$ and hence are done by (II.1.1).

(b) will be a simple consequence of Theorem II.7.2 below, the proof of which will use (a). See also Section II.10 of Harris (1963). ∎

These results suggest we consider a sequence of critical Galton-Watson branching processes $\{Z_0^{(n)} : n \in \mathbb{N}\}$ as above but with initial conditions $Z_0^{(n)}$ satisfying $Z_0^{(n)}/n \to x$, and define $X_t^{(n)} = Z_{[nt]}^{(n)}/n$. Indeed it is an easy exercise to see from the above that $X_1^{(n)}$ converges weakly to a Poisson sum of independent exponential masses.

Notation. E denotes a Polish space. Let $D(E) = D(\mathbb{R}_+, E)$ be the Polish space of cadlag paths from $\mathbb{R}_+$ to E with the Skorokhod J_1-topology. Let $C(E) = C(\mathbb{R}_+, E)$ be the Polish space of continuous E-valued paths with the topology of uniform convergence on compacts. Let $Y_t(y) = y(t)$ for $y \in D(E)$.

Theorem II.1.2. (Feller (1939, 1951)) $X^{(n)} \overset{w}{\Rightarrow} X$ in $D(\mathbb{R})$, where X is the unique solution of

$$(FE) \qquad X_t = x + \int_0^t \sqrt{\gamma X_s}\, dB_s,$$

where B is a one-dimensional Brownian motion.

Proof. We will prove much more below in Theorem II.5.1. The uniqueness holds by Yamada-Watanabe (1971). ∎

We call the above process Feller's branching diffusion with parameter γ.

2. Spatial Motion

We now give our branching population a spatial structure. Individuals are "located" at a point in a Polish space E. This structure will also usually allow us to trace back the genealogy of individuals in the population.

Notation. $\mathcal{E}$ = Borel subsets of $E \equiv \mathcal{B}(E)$,

$C_b(E) = \{f : E \to \mathbb{R} : f \text{ bounded and continuous}\}$ with the supnorm, $\|\ \|$,

$\mathcal{D} = \mathcal{B}\big(D(E)\big)$, $(\mathcal{D}_t)_{t\ge 0}$ is the canonical right-continuous filtration on $D(E)$,

$\mu(f) = \int f d\mu$ for a measure μ and integrable function f.

Assume

$$(II.2.1) \qquad Y = (D, \mathcal{D}, \mathcal{D}_t, Y_t, P^x) \text{ is a Borel strong Markov process (BSMP)}$$

with semigroup $P_t f(x) = P^x(f(Y_t))$. "Borel" means $x \to P^x(A)$ is $\mathcal{E}$-measurable for all $A \in \mathcal{D}$. The other required properties here are $Y_0 = x$ P^x-a.s. and the strong Markov property. Evidently our BSMP's have cadlag paths. These assumptions are either much too restrictive or far too abstract, depending on your upbringing. At the risk of offending one of these groups we impose an additional condition:

$$(II.2.2) \qquad P_t : C_b(E) \to C_b(E).$$

This is only needed to facilitate our construction of Dawson-Watanabe superprocesses as limits of branching particle systems and keep fine topologies and Ray compactifications at bay.

Standard arguments (see Exercise II.2.1 below or the proof of Theorem I.9.21 in Sharpe (1988)) show that (II.2.2) implies

Y is a Hunt process, i.e., if $\{T_n\}$ are $\{\mathcal{D}_t\}$-stopping times such that

(QLC) $T_n \uparrow T < \infty$ $P^x a.s.$, then $Y(T_n) \to Y(T)$ $P^x - a.s.$

In particular, $Y_t = Y_{t-}$ $P^x - a.s.$ for all $t > 0$.

Definition. $\phi \in \mathcal{D}(A)$ iff $\phi \in C_b(E)$ and for some $\psi \in C_b(E)$,

$$\phi(Y_t) - \phi(Y_0) - \int_0^t \psi(Y_s)ds \text{ is a } P^x\text{-martingale for all } x \text{ in } E.$$

It is easy to see ψ is unique if it exists and so we write $\psi = A\phi$ for $\phi \in \mathcal{D}(A)$.

Notation. $\overset{bp}{\to}$ denotes bounded pointwise convergence.

Proposition II.2.1. $\phi \in \mathcal{D}(A) \Leftrightarrow \phi \in C_b(E)$ and $\frac{P_t\phi-\phi}{t}\overset{bp}{\to}\psi$ as $t \downarrow 0$ for some $\psi \in C_b(E)$. In this case, $\psi = A\phi$ and for any $s \geq 0$, $P_s\phi \in \mathcal{D}(A)$ and

$$(II.2.3) \qquad AP_s\phi = P_sA\phi = \frac{\partial}{\partial s}P_s\phi.$$

Proof. ($\Leftarrow$) If $s \geq 0$, our assumption and the semigroup property show that

$$\frac{P_{s+t}\phi - P_s\phi}{t}\overset{bp}{\to}P_s\psi \quad \text{as} \quad t\downarrow 0 \quad \forall s \geq 0.$$

(QLC) implies $P_s\psi(x)$ is continuous in s for each x. An easy calculus exercise shows that a continuous function with a continuous right derivative is differentiable, and so from the above we have

$$(II.2.4) \qquad \frac{\partial}{\partial s}P_s\phi = P_s\psi,$$

and so

$$P^x\left(\phi(Y_t) - \phi(Y_0) - \int_0^t \psi(Y_s)ds\right) = P_t\phi(x) - \phi(x) - \int_0^t P_s\psi(x)ds = 0.$$

The Markov property now shows the process in the above expectation is a martingale and so $\phi \in \mathcal{D}(A)$ with $A\phi = \psi$.

($\Rightarrow$) Let $\phi \in \mathcal{D}(A)$ and $s \geq 0$.

$$\big(P_{t+s}\phi(x) - P_s\phi(x)\big)/t = P^x\left(P^{Y_s}\left(\int_0^t A\phi(Y_r)dr/t\right)\right)\overset{bp}{\to}P_s(A\phi)(x) \quad \text{as} \quad t\downarrow 0,$$

where the last limit holds by Dominated Convergence. Taking $s = 0$, one completes the proof of ($\Rightarrow$). For $s > 0$, one may use the above argument and ($\Leftarrow$) to see $P_s\phi \in \mathcal{D}(A)$ and get the first equality in (II.2.3). The second equality follows from (II.2.4) with $\psi = A\phi$. ∎

Let $U_\lambda f(x) = E^x\big(\int_0^\infty e^{-\lambda t}f(Y_t)dt\big)$ denote the λ-resolvent of Y for $\lambda > 0$. Clearly $U_\lambda : C_b(E) \to C_b(E)$ by (II.2.2).

Proposition II.2.2.
(a) $\forall\phi \in \mathcal{D}(A) \qquad U_\lambda(\lambda - A)\phi = \phi$.
(b) $\forall\phi \in C_b(E) \qquad U_\lambda\phi \in \mathcal{D}(A) \quad \text{and} \quad (\lambda - A)U_\lambda\phi = \phi$.

Proof. (a)

$$\begin{aligned}
U_\lambda A\phi(x) &= \int_0^\infty e^{-\lambda t} P_t A\phi(x)dt \\
&= \int_0^\infty e^{-\lambda t} \frac{\partial}{\partial t}\big(P_t\phi(x)\big)dt \quad \text{(by II.2.3)} \\
&= e^{-\lambda t} P_t\phi(x)\Big|_{t=0}^{t=\infty} + \lambda \int_0^\infty e^{-\lambda t} P_t\phi(x)dt \\
&= -\phi(x) + \lambda U_\lambda \phi(x).
\end{aligned}$$

(b)

$$(II.2.5) \qquad \begin{aligned}
U_\lambda\phi(Y_t) &= E^x\Big(\int_t^\infty e^{-\lambda u}\phi(Y_u)du\Big|\mathcal{D}_t\Big)e^{\lambda t} \\
&= e^{\lambda t}\Big[M_t - \int_0^t e^{-\lambda u}\phi(Y_u)du\Big],
\end{aligned}$$

where M_t denotes the martingale $E^x\big(\int_0^\infty e^{-\lambda u}\phi(Y_u)du \mid \mathcal{D}_t\big)$. Some stochastic calculus shows that ($\overset{m}{=}$ means equal up to martingales)

$$U_\lambda\phi(Y_t)\overset{m}{=}\int_0^t \lambda e^{\lambda s}\left(M_s - \int_0^s e^{-\lambda u}\phi(Y_u)du\right)ds - \int_0^t \phi(Y_s)ds = \int_0^t \lambda U_\lambda\phi(Y_s) - \phi(Y_s)ds,$$

where we have used (II.2.5). This implies $U_\lambda\phi \in \mathcal{D}(A)$ and $AU_\lambda\phi = \lambda U_\lambda\phi - \phi$. ∎

Notation. $b\mathcal{E}$ (respectively, $\mathcal{E}_+$) is the set of bounded (respectively, non-negative) $\mathcal{E}$-measurable functions. If $\mathcal{H} \subset b\mathcal{E}$, $\overline{\mathcal{H}}^{bp}$ is the smallest set containing $\mathcal{H}$ and closed under $\overset{bp}{\to}$, and $\mathcal{H}_+$ is the set of non-negative functions in $\mathcal{H}$.

Corollary II.2.3. $\overline{\mathcal{D}(A)}^{bp} = b\mathcal{E}, \quad \overline{(\mathcal{D}(A)_+)}^{bp} = b\mathcal{E}_+$.

Proof. If $\phi \in C_b(E)$, $P_t\phi\overset{bp}{\to}\phi$ as $t \downarrow 0$ and so it follows that $\lambda U_\lambda\phi\overset{bp}{\to}\phi$ as $\lambda \to \infty$, and so $\phi \in \overline{\mathcal{D}(A)}^{bp}$. The result follows trivially. ∎

Exercise II.2.1. Prove that Y satisfies (QLC).
Hint. (Following the proof of Theorem I.9.21 of Sharpe (1988).) Let

$$X = \lim Y(T_n) \in \{Y(T-), Y(T)\}.$$

It suffices to consider T bounded and show $E^x(g(X)h(Y_T)) = E^x(g(X)h(X))$ for all $g, h \in C_b(E)$ (why?). As in the proof of Corollary II.2.3 it suffices to consider $h = U_\lambda f$, where $f \in C_b(E)$ and $\lambda > 0$. Complete the required argument by using the strong Markov property of Y and the continuity of $U_\lambda f$.

Here are some possible choices of Y.

Examples II.2.4. (a) $Y_t \in \mathbb{R}^d$ is d-dimensional Brownian motion.

$$C_b^2(\mathbb{R}^d) = \{\phi : \mathbb{R}^d \to \mathbb{R}, \phi \text{ is } C^2 \text{ with bounded partials of order 2 or less}\} \subset \mathcal{D}(A)$$

and $A = \frac{\Delta\phi}{2}$ for $\phi \in C_b^2(\mathbb{R}^d)$ by Itô's Lemma. In this case we will write $\mathcal{D}(\Delta/2)$ for $\mathcal{D}(A)$.
(b) $Y_t \in \mathbb{R}^d$ is the d-dimensional symmetric stable process of index $\alpha \in (0,2)$ scaled so that $P^x\left(e^{i\theta\cdot Y_t}\right) = e^{i\theta\cdot x - t|\theta|^\alpha}$, where $|y|$ is the Euclidean norm of y. If $\nu(dy) = c|y|^{-d-\alpha}dy$ for an appropriate $c > 0$, then for $\phi \in C_b^2(\mathbb{R}^d) \subset \mathcal{D}(A)$

$$A\phi(x) = \int \left[\phi(x+y) - \phi(x) - \vec{\nabla}\phi(x) \cdot \frac{y}{1+|y|^2}\right] \nu(dy),$$

as can be easily seen, e.g., from the stochastic calculus for point processes in Ikeda-Watanabe (1981, p. 65–67) (see also Revuz-Yor (1990, p. 263)).

In both the above examples

$$C_b^\infty(\mathbb{R}^d) = \{\phi \in C_b(\mathbb{R}^d) : \text{all partial derivatives of } \phi \text{ are in } C_b(\mathbb{R}^d)\}$$

is a core for A in that the *bp*-closure of $\{(\phi, A\phi) : \phi \in C_b^\infty(\mathbb{R}^d)\}$ contains $\{(\phi, A\phi) : \phi \in \mathcal{D}(A)\}$. To see this first note that if $\phi \in \mathcal{D}(A)$ has compact support, then

$$P_t\phi(x) = \int p_t(y-x)\phi(y)dy \in C_b^\infty(\mathbb{R}^d) \text{ for } t > 0$$

because Y has a transition density, p_t, all of whose spatial derivatives are bounded and continuous. In the stable case the latter is clear from Fourier inversion because

$$|\theta|^m \int p_t(y-x)e^{i\theta\cdot y}dy = |\theta|^m e^{i\theta\cdot x - t|\theta|^\alpha}$$

is bounded and integrable in θ for all $m \in \mathbb{N}$. Now choose $\{\psi_n\} \subset C_b^\infty$ with compact support so that $\psi_n \uparrow 1$ and

$$\{|x| \le n\} \subset \{\psi_n = 1\} \subset \{\psi_n > 0\} \subset \{|x| < n+1\}.$$

If $\phi \in \mathcal{D}(A)$ and $\phi_n = \phi\psi_n$, then an integration by parts shows that $\phi_n \in \mathcal{D}(A)$ and $A\phi_n = \psi_n A\phi + \phi A\psi_n$. The above shows that $P_{1/n}\phi_n \in C_b^\infty$. Dominated Convergence implies that $P_{1/n}\phi_n \xrightarrow{bp} \phi$, and (II.2.3) and a short calculation shows that

$$AP_{1/n}\phi_n = P_{1/n}A\phi_n = P_{1/n}(\psi_n A\phi + \phi A\psi_n) \xrightarrow{bp} A\phi.$$

This proves the above claim.

Notation. $M_1(E)$ is the space of probabilities on a Polish space E and its Borel σ-field, equipped with the topology of weak convergence. $C_K(E)$ is the space of continuous function on E with compact support, equipped with the topology of uniform convergence.

Exercise II.2.2. Assume $Y_t \in \mathbb{R}^d$ is d-dimensional Brownian motion with $d > 2$ and $U_0 f$ is defined as above but with $\lambda = 0$.

(a) Show that if $f \geq 0$ is Borel measurable on $\mathbb{R}^d$, then as $\lambda \downarrow 0$, $U_\lambda f(x) \uparrow U_0 f(x) = k_d \int |y-x|^{2-d} f(y) ds \leq \infty$, for some $k_d > 0$.

(b) Show that $U_0 : C_K(\mathbb{R}^d) \to \mathcal{D}(A)$, $AU_0\phi = -\phi$ for all $\phi \in C_K(\mathbb{R}^d)$, and $U_0 A\phi = -\phi$ for all $\phi \in C_K(\mathbb{R}^d) \cap \mathcal{D}(A)$.

Hint. One approach to (b) is to show that for $\phi \in C_K(\mathbb{R}^d)$, as $\lambda \downarrow 0$, $U_\lambda \phi \stackrel{bp}{\to} U_0\phi \in C_b$ and $AU_\lambda \phi \stackrel{bp}{\to} -\phi$.

Example II.2.4. (c) Suppose, in addition, that our BSMP Y satisfies

$$(PC) \qquad\qquad x \to P^x \quad \text{is continuous from } E \text{ to } M_1\big(D(E)\big).$$

This clearly implies (II.2.2).

Under (PC) we claim that the previous hypotheses are satisfied by the path-valued process $t \to (t, Y^t) \equiv \big(t, Y(\cdot \wedge t)\big) \in \mathbb{R}_+ \times D(E)$. To be more precise, let $\hat{E} = \{(t, y^t) : t \geq 0, y \in D(E)\}$ with the subspace topology it inherits from $\mathbb{R}_+ \times D(E)$, $\hat{\mathcal{E}} = \mathcal{B}(\hat{E})$, and if $y, w \in D(E)$ and $s \geq 0$, let

$$(y/s/w)(t) = \begin{cases} y(t) & t < s \\ w(t-s) & t \geq s \end{cases} \qquad \Big(\in D(E)\Big).$$

Note that $\hat{E}$ is Polish as it is a closed subset of the Polish space $\mathbb{R}_+ \times D(E)$.

Definition. Let $W_t : D(\hat{E}) \to \hat{E}$ denote the coordinate maps and for $(s, y) \in \hat{E}$, define $\hat{P}_{s,y}$ on $D(\hat{E})$ with its Borel σ-field, $\hat{\mathcal{D}}$, by

$$\hat{P}_{s,y}(W_\cdot \in A) = P^{y(s)}\big((s + \cdot, y/s/Y^\cdot) \in A\big),$$

i.e., under $\hat{P}_{s,y}$ we run y up to time s and then tag on a copy of Y starting at $y(s)$.

Proposition II.2.5. $(W, (\hat{P}_{s,y})_{(s,y)\in\hat{E}})$ is a BSMP with semigroup

$$\hat{P}_t : C_b(\hat{E}) \to C_b(\hat{E}).$$

Proof. This is a routine if somewhat tedious exercise. Let $(\hat{\mathcal{D}}_t)$ be the canonical right-continuous filtration on $D(\hat{E})$. Fix $u \geq 0$ and to check the Markov property at time u, let $(s, y) \in \hat{E}$, $A \in \hat{\mathcal{D}}_u$, ψ be a bounded measurable function on $\hat{E}$ and $T \geq u$. Also set

$$\tilde{A} = \{w \in D(E) : (v \to (s + v, (y/s/w^v)) \in A\} \in \mathcal{D}_u,$$

and

$$\tilde{\psi}(w) = \psi(s + u + t, y/s/(w^{u+t})),\ w \in D(E).$$

Then

$$\int 1_A \psi(W_{u+t}) d\hat{P}_{s,y} = \int 1_{\tilde{A}}(Y)\tilde{\psi}(Y^{u+t}) dP^{y(s)}$$
$$= \int 1_{\tilde{A}}(Y(\omega)) P^{Y_u(\omega)}(\tilde{\psi}(Y^u(\omega)/u/Y^t) dP^{y(s)}(\omega) \quad \text{(by the Markov property for } Y)$$
$$= \int 1_{\tilde{A}}(Y(\omega)) P^{Y_u(\omega)}(\psi(s+u+t, (y/s/Y^u(\omega))/s+u/Y^t))\, dP^{y(s)}(\omega)$$
$$= \int 1_A(s+\cdot, y/s/Y^{\cdot}(\omega)) \hat{P}_{s+u, y/s/Y^u(\omega)}(\psi(W_t)) dP^{y(s)}(\omega)$$
$$= \int 1_A \hat{P}_{W_u(\omega)}(\psi(W_t)) d\hat{P}_{s,y}(\omega).$$

This proves the Markov property at time u.

Turning now to the semigroup $\hat{P}_t$, let $f \in C_b(\hat{E})$ and suppose $(s_n, y_n) \to (s_\infty, y_\infty)$ in $\hat{E}$. Note that if $T > \sup_n s_n$ is a continuity point of y_∞, then

$$y_n(s_n) = y_n(T) \to y_\infty(T) = y_\infty(s_\infty).$$

Therefore by (PC) and Skorohod's representation (see Theorem 3.1.8 of Ethier and Kurtz (1986)) we may construct a sequence of processes, $\{Y_n : n \in \mathbb{N} \cup \{\infty\}\}$ such that Y_n has law $P^{y_n(s_n)}$ and $\lim Y_n = Y_\infty$ in $D(\hat{E})$ a.s. Now use the fact that

$$\begin{aligned} &s_n \to s_\infty,\ (y_n, Y_n) \to (y_\infty, Y_\infty) \text{ in } D(E)^2, \text{ and } Y_n(0) = y_n(s_n) \\ &\text{imply} \quad y_n/s_n/Y_n \to y_\infty/s_\infty/Y_\infty \text{ as } n \to \infty \text{ in } D(E). \end{aligned}$$

This is a standard exercise in the Skorokhod topology on $D(E)$ which is best left to the reader. Note that the only issue is the convergence near s_∞ and here the condition $Y_n(0) = y_n(s_n)$ avoids the possibility of having distinct jump times approach s_∞ in the limit. The above implies that $\lim_{n\to\infty} f(y_n/s_n/Y_n^t) = f(y_\infty/s_\infty/Y_\infty^t)$ a.s. and therefore

$$\lim_{n\to\infty} \hat{P}_t f(s_n, y_n) = \lim_{n\to\infty} E(f(y_n/s_n/Y_n^t)) = E(f(y_\infty/s_\infty/Y_\infty^t)) = \hat{P}_t f(s_\infty, y_\infty).$$

This shows that $\hat{P}_t : C_b(\hat{E}) \to C_b(\hat{E})$. The strong Markov property of W now follows from this and the ordinary Markov property by a standard approximation of a stopping time by a sequence of countably-valued stopping times. Also the Borel measurability of $(s, y) \to \hat{P}_{s,y}(\psi(W))$ is now clear for ψ a bounded continuous function of finitely many coordinates, in which case this function is continuous by the above, and hence for all bounded and measurable ψ by the usual bootstrapping argument. Finally $\hat{P}_{s,y}(W(0) = (s, y)) = 1$ is clear from the definitions. ∎

Exercise II.2.3. Let Y_t be a Feller process with a strongly continuous semigroup $P_t : C_\ell(E) \to C_\ell(E)$, where $C_\ell(E)$ is the space of continuous functions on a locally compact metric space (E, d) with finite limit at ∞. Show that (PC) holds.

Hint. Let $x_n \to x$. It suffices to show $\{P^{x_n}\}$ is tight on $D(E)$ (why?). By Aldous' criterion (Walsh (1986), Thm. 6.8(a)) it suffices to consider a sequence of stopping times $\{T_n\}$, $T_n \le t_0 < \infty$, a sequence $\delta_n \downarrow 0$, and show

$$\lim_{n\to\infty} P^{x_n}\big(d(Y_{T_n+\delta_n}, Y_{T_n}) > \varepsilon\big) = 0 \qquad \forall \varepsilon > 0.$$

Exercise II.2.4. Give an example of a Markov process satisfying the hypotheses (II.2.1) and (II.2.2) of this Section but for which (PC) fails.

Hint. Take $E = \mathbb{R}_+$ and consider the generator

$$Af(y) = (f(1)-f(0))1(y=0)+(f(1/y)-f(y))1(0<y\le 1)+y(f(1)-f(y))1(y>1).$$

Thanks go to Tom Kurtz for suggesting this example.

3. Branching Particle Systems

Let Y be the E-valued Markov process from the previous Section and introduce a drift function $g \in C_b(E)$ and branching variance (or rate) function $\gamma \in C_b(E)_+$. Recall that G_+ denotes the non-negative elements in the set G. The state space for our limiting and approximating processes will be $M_F(E)$, the space of finite measures on E with the topology of weak convergence, and we choose an initial state $X_0 \in M_F(E)$.

For $N \in \mathbb{N}$ and $x \in E$, $\nu^N(x,\cdot) \in M_1(\mathbb{Z}_+)$ is the offspring law for a parent located at x. We assume $x \to \nu^N(x,\cdot)$ is measurable and satisfies ($\overset{ucb}{\to}$ denotes convergence on E which is uniform on compacts and bounded on E):

$$(II.3.1)\qquad \begin{aligned} &\text{(a)} \int k\nu^N(x,dk) = 1 + \frac{g_N(x)}{N} \quad \text{where} \quad g_N \overset{ucb}{\to} g \text{ as } N\to\infty,\\ &\text{(b) Var } \big(\nu^N(x,\cdot)\big) = \gamma_N(x) \quad \text{where} \quad \gamma_N \overset{ucb}{\to}\gamma \text{ as } N\to\infty,\\ &\text{(c) } \exists\, \delta>0 \quad \text{such that} \quad \sup_{N,x}\int k^{2+\delta}\nu^N(x,dk) < \infty. \end{aligned}$$

Remarks II.3.1. (1) At the cost of complicating our arguments somewhat, (II.3.1) (c) could be weakened to the uniform integrability of k^2 with respect to $\{\nu^N(x,\cdot) : x \in E, N \in \mathbb{N}\}$.

(2) Given $\gamma \in C_b(E)_+$ and $g \in C_b(E)$ it is not hard to see there is a sequence $\{\nu^N\}$ satisfying (II.3.1). In fact there is a $k \in \mathbb{Z}$, $k \ge 2$ and functions p_N and q_N such that

$$\nu^N(x,\cdot) = \delta_0(1-p_N(x)-q_N(x)) + \delta_1 p_N(x) + \delta_k q_N(x)$$

will satisfy (II.3.1) with $g_N \equiv g$ for $N \ge N_0$.

Exercise II.3.1. Prove this.

Hint. A simple calculation shows that if (II.3.1(a,b)) hold for ν^N as above with $g_N = g$, then $p_N(x) = 1 + g(x)/N - \alpha_N(x)/(k-1)$ and $q_N(x) = \alpha_N(x)/(k^2-k)$, where

$$\alpha_N(x) = \gamma_N(x) + g(x)/N + g(x)^2/N^2.$$

Let $\eta_N = \|g\|_\infty/N + \|g\|_\infty^2/N^2 + \|g\|_\infty/\sqrt{N}$ and set $\gamma_N(x) = \gamma(x) \vee \eta_N$. Show that you can choose k sufficiently large so that $p_N(x), q_N(x) \ge 0$, $p_N(x)+q_N(x) \le 1$ for N large, and (II.3.1) is valid with $g_N = g$ for such an N.

We now describe a system of branching particles which undergo near critical branching at times k/N according to the laws $\nu_N(x,\cdot)$ where x is the location of the parent. In between branch times particles migrate as independent copies of the process Y from the previous section. It will be important to have a labeling scheme to refer to the branches of the resulting tree of Y-processes. We follow the arboreal labeling of Walsh (1986) – in fact this section is really the missing Ch. 10 of Walsh's SPDE notes. We have decided to work in a discrete time setting but could just as well work in the continuous time setting in which inter-branch intervals are exponential rate N random variables.

We label particles by multi-indices

$$\alpha\in I=\bigcup_{n=0}^{\infty}\mathbb{N}^{\{0,\dots,n\}}=\{(\alpha_0,\dots,\alpha_n):\alpha_i\in\mathbb{N},n\in\mathbb{Z}_+\}.$$

Let $|(\alpha_0,\dots,\alpha_n)|=n$ be the generation of α and write

$$\beta<\alpha\Leftrightarrow\beta=(\alpha_0,\dots,\alpha_i)\equiv\alpha|i\text{ for some }i\le|\alpha|,$$

i.e. if β is an ancestor of α. We let $\alpha\vee k=(\alpha_0,\dots\alpha_n,k)$ denote the k^{th} offspring of α and $\alpha\wedge\beta$ denote the "greatest common ancestor" of α and β (set $\alpha\wedge\beta=\phi$ if $\alpha_0\ne\beta_0$ and $|\phi|=-\infty$), and let $\pi\alpha=(\alpha_0,\dots,\alpha_{n-1})$ denote the parent of α if $n>0$.

Adjoin Δ to E as a discrete point to form $E_\Delta=E\cup\{\Delta\}$ and let P^Δ be point mass at the constant path identical to Δ. Let $\Omega=E_\Delta^{\mathbb{N}}\times D(E_\Delta)^{\mathbb{N}}\times\mathbb{Z}_+^{\mathbb{N}}$ and let $\mathcal{F}$ denote its product σ-field. Sample points in Ω are denoted by

$$\omega=((x_i,i\in\mathbb{N}),(Y^\alpha,\alpha\in I),(N^\alpha,\alpha\in I)).$$

Now fix $N\in\mathbb{N}$ and define a probability $P=P^N$ on $(\Omega,\mathcal{F})$ as follows:
$(II.3.2)$

(a) $(x_i,i\le M_N)$ is a Poisson point process with intensity $NX_0(\cdot)$ and $x_i=\Delta$ if $i>M_N$.

(b) Given $\mathcal{G}_n=\sigma(x_i,i\in\mathbb{N})\vee\sigma(N^\beta,Y^\beta,|\beta|<n)$, $\{Y^\alpha:|\alpha|=n\}$ are (conditionally) independent and (for $|\alpha|=n$)

$$P(Y^\alpha\in A|\mathcal{G}_n)(\omega)=P^{x_{\alpha_0}(\omega)}\Big(Y\Big(\cdot\wedge\Big(\frac{n+1}{N}\Big)\Big)\in A\Big|Y\Big(\cdot\wedge\frac{n}{N}\Big)=Y^{\pi\alpha}_{\cdot}(\omega)\Big)$$

where $Y^{\pi\alpha}\equiv x_{\alpha_0}$ if $|\alpha|=0$. That is, $Y^\alpha\Big(\cdot\wedge\dfrac{|\alpha|}{N}\Big)=Y^{\pi\alpha}(\cdot)$and given $\mathcal{G}_n$, $\{Y^\alpha|_{[|\alpha|/N,(|\alpha|+1)/N]}:|\alpha|=n\}$ evolve as independent copies of Y starting from $Y^{\pi\alpha}(|\alpha|/N)$, and stopped at $(|\alpha|+1)/N$.

(c) Given $\bar{\mathcal{G}}_n=\mathcal{G}_n\vee\sigma(Y^\alpha:|\alpha|=n)$, $\{N^\alpha:|\alpha|=n\}$ are (conditionally) independent and $P(N^\alpha\in\cdot|\bar{\mathcal{G}}_n)(\omega)=\nu^N(Y^\alpha((|\alpha|+1)/N,\omega),\cdot)$.

It should be clear from the above that $\{Y^\alpha:\alpha\in I\}$ is an infinite tree of branching Y processes, where $Y_t^\alpha=Y_t^\beta$ for $0\le t<(|\alpha\wedge\beta|+1)/N$. Let $\underline{t}=[Nt]/N$ for $t\ge0$ where $[x]$ is the integer part of x, set $T=T_N=\{kN^{-1}:k\in\mathbb{Z}_+\}$ and let $\tau = 1/N$. It will be convenient to work with respect to the right continuous filtration given by

$$\mathcal{F}_t = \mathcal{F}_t^N = \sigma\left((x_i)_{i\in\mathbb{N}}, (Y^\alpha, N^\alpha)_{|\alpha|<N\underline{t}}\right) \vee \left(\bigcap_{r>t} \sigma\left(Y_s^\alpha : |\alpha| = N\underline{t}, s \le r\right)\right)$$

It also will be useful to introduce the slightly larger σ-field

$$\overline{\mathcal{F}}_{\underline{t}} = \mathcal{F}_{\underline{t}} \vee \sigma\big(Y^\alpha : |\alpha| = N\underline{t}\big).$$

Here are some consequences of our definition of P for each $\alpha \in I$ and $\underline{t} = |\alpha|/N$:

$(II.3.3)$ $\{(Y^\alpha, \mathcal{F}_s) : s \ge \underline{t}\}$ is a Markov process,and for all $s \in [\underline{t}, \underline{t}+\tau]$,
$P(Y^\alpha(s+\cdot) \in A|\mathcal{F}_s)(\omega) = P^{Y_s^\alpha(\omega)}(Y(\cdot \wedge (\underline{t}+\tau-s)) \in A)$ a.s.
for all $A \in \mathcal{D}$.

$(II.3.4)$ $P(Y^\alpha \in A|\mathcal{F}_0)(\omega) = P^{x_{\alpha_0}(\omega)}(Y(\cdot \wedge (\underline{t}+\tau)) \in A)$ a.s. for all $A \in \mathcal{D}$.

$(II.3.5)$ $$P(N^\alpha \in \cdot|\bar{\mathcal{F}}_{\underline{t}}) = \nu^N(Y^\alpha_{\underline{t}+\tau}, \cdot) \text{ a.s.}$$

Clearly (II.3.5) is a restatement of (II.3.2)(c). (II.3.4) should be clear from (II.3.3) and (II.3.2)(b) (one can, for example, induct on $|\alpha|$). To prove (II.3.3), it suffices to prove the stated equality for $s \in [\underline{t}, \underline{t}+\tau)$, so fix such an s. The stated result is an easy consequence of (II.3.2)(b) if $\mathcal{F}_s$ is replaced by the smaller σ-field $\mathcal{G}_{N\underline{t}} \vee \mathcal{F}_s^{Y^\alpha}$, where $\mathcal{F}_s^{Y^\alpha}$ is the right continuous filtration generated by Y^α. Now use the fact that $\mathcal{H}^\alpha \equiv \sigma(Y^\beta : |\beta| = N\underline{t}, \beta \ne \alpha)$ is conditionally independent of Y^α given $\mathcal{G}_{N\underline{t}}$ (by (II.3.2(b)) to see that the stated result is valid if $\mathcal{F}_s$ is replaced by the larger σ-field $\mathcal{G}_{N\underline{t}} \vee \mathcal{F}_s^{Y^\alpha} \vee \mathcal{H}^\alpha$. Now condition this last equality with respect to $\mathcal{F}_s$ to obtain (II.3.3).

Remark II.3.2. If $X_0^N = \frac{1}{N}\sum_1^{M_N} \delta_{x_i}$, an easy consequence of (II.3.2)(a) and the weak law of large numbers is

$(II.3.6)$ $$X_0^N(\phi) \xrightarrow{P} X_0(\phi) \text{ as } N \to \infty \text{ for any bounded measurable } \phi \text{ on } E.$$

Note also that $E(X_0^N(\cdot)) = X_0(\cdot)$. From (II.3.6) it is easy to show that $X_0^N \xrightarrow{P} X_0$. For example, one could use the existence of a countable convergence determining class of functions on E (see the proof of Theorem 3.4.4 in Ethier and Kurtz (1986)).

Instead of assuming $\{x_i^N, i \le M_N\}$ is as in (II.3.2)(a), we could assume more generally that $\{x_i^N, i \le M_N\}$ are random points (M_N is also random) chosen so that $X_0^N = \frac{1}{N}\sum_1^{M_N} \delta_{x_i^N} \xrightarrow{P} X_0$, $\sup_N E(X_0^N(1)^2) < \infty$, and $E(X_0^N(\cdot)) \le c_0 X_0(\cdot)$ as measures. The only required change is to include c_0 as multiplicative factor in the upper bound in Lemma II.3.3 below. For example, we could assume $\{x_i, i \le M_N\}$ are i.i.d. with law $X_0(\cdot)/X_0(1)$ and $M_N = [NX_0(1)]$ and set $c_0 = 1$.

The interpretation of (II.3.2) (b,c) is that for $|\alpha| = N\underline{t}$, the individual labelled by α follows the trajectory Y^α on $[\underline{t}, \underline{t}+\tau]$ and at time $\underline{t}+\tau$ dies and is replaced by its N^α children. The next step is to use the N^α to prune the infinite tree of branching Y^α processes. The termination time of the α^{th} branch is

$$\zeta^\alpha = \begin{cases} 0, \quad \text{if} \quad x_{\alpha_0} = \Delta \\ \min\Big\{(i+1)/N : i < |\alpha|, N^{\alpha|i} < \alpha_{i+1}\Big\}, \quad \text{if this set is not } \emptyset \text{ and } x_{\alpha_0} \neq \Delta \\ (|\alpha|+1)/N, \quad \text{otherwise.} \end{cases}$$

Note in the first case, the α^{th} particle was never born since $\alpha_0 > M_N$. In the second case, $\alpha_{i+1} > N^{\alpha|i}$ means the α_{i+1}st offspring of $\alpha|i$ doesn't exist. Finally in the last instance, the family tree of α is still alive at $(|\alpha|+1)/N$ but we have run out of data to describe its state beyond this final time.

We write $\alpha \overset{\omega}{\sim} t$ (or $\alpha \sim t$) iff $|\alpha|/N \le t < (|\alpha|+1)/N = \zeta^\alpha$, i.e., iff α labels a particle alive at time t. Clearly $\alpha \sim t$ iff $\alpha \sim \underline{t}$. Note that we associate α with the particle alive on $[|\alpha|/N, (|\alpha|+1)/N]$, although of course Y^α_s, $s < (|\alpha|+1)/N$ describes the past history of its ancestors. From Feller's theorem (Theorem II.1.2) it is natural to assign mass $\frac{1}{N}$ to each particle alive at time t and define

$$X^N_t = \frac{1}{N}\sum_{\alpha\sim t}\delta_{Y^\alpha_t} = \frac{1}{N}\sum_{\alpha\sim \underline{t}}\delta_{Y^\alpha_t}, \quad \text{i.e.,} \quad X^N_t(A) = \#\{Y^\alpha_t \in A : \alpha \sim t\}/N, \quad A \in \mathcal{E}.$$

Since $N^\alpha < \infty$ for all $\alpha \in I$ a.s., clearly $X^N_t \in M_F(E)$ for all $t \ge 0$ a.s. Note also that $Y^\alpha \in D(E)$ for all α with $x_{\alpha_0} \neq \Delta$ a.s., and therefore $X^N_t = \frac{1}{N}\sum_{\alpha\sim\underline{t}}\delta_{Y^\alpha_t}$ has sample paths in $D\big(M_F(E)\big)$ a.s. on each $[\underline{t}, \underline{t}+\tau)$, and hence on all of $\mathbb{R}_+$. The associated *historical process* is

$$H^N_t = \frac{1}{N}\sum_{\alpha\sim t}\delta_{Y^\alpha_{\cdot\wedge t}} \in M_F\big(D(E)\big).$$

Again $H^N_\cdot \in D\big(\mathbb{R}_+, M_F(D(E))\big)$. Therefore X^N_t is the (normalized) empirical measure of the particles alive at time t while H^N_t records the past trajectories of the ancestors of particles alive at time t. Clearly we have

$$X^N_t(\phi) = \int \phi(y_t) H^N_t(dy).$$

Exercise II.3.2. Show that

(i) $\{\alpha \sim t\} \in \mathcal{F}_t$.

(ii) X^N_t is $\mathcal{F}_t$-measurable.

(A trivial exercise designed only to convince you that $|\alpha| < N\underline{t}$ in the above definition is correct.)

Our goal is to show $X^N \overset{w}{\Rightarrow} X$ in $D\big(M_F(E)\big)$ and characterize X as the unique solution of a martingale problem. The weak convergence of H^N to the associated historical process H will then follow easily by considering the special case in Example II.2.4 (c). As any student of Ethier and Kurtz (1986) knows, the proof proceeds in two steps:

1. Tightness of $\{X^N\}$ and derivation of limiting martingale problem.
2. Uniqueness of solutions to the martingale problem.

These are carried out in the next 3 sections.

We close this section with a simple bound for the first moments.

Notation. $P^\mu = \int P^x(\cdot)\mu(dx), \mu \in M_F(E)$. If $y \in C(E)$, let $y^t(s) = y(t \wedge s)$.

Lemma II.3.3. Let $g_\infty = \sup_N \|g_N\|_\infty$ and $\overline{g} = \sup\{g_N(x) : x \in E, N \in \mathbb{N}\}$.

(a) If $\psi : D(\mathbb{R}_+, E) \to \mathbb{R}_+$ is Borel, then for any $t \geq 0$

$$E\left(H_t^N(\psi)\right) \leq e^{\overline{g}\underline{t}} E^{X_0}\left(\psi(Y^t)\right) \leq e^{g_\infty \underline{t}} E^{X_0}\left(\psi(Y^t)\right).$$

In particular, $E\big(X_t^N(\phi)\big) \leq e^{\overline{g}t} E^{X_0}\big(\phi(Y_t)\big)\ \forall \phi \in \mathcal{E}_+$.

(b) For all $x, K > 0$ and for all $N \geq N_0(g_\infty)$,

$$P\Big(\sup_{t \leq K} X_t^N(1) \geq x\Big) \leq e^{3g_\infty K} X_0(1) x^{-1}.$$

Proof. (a) Let $P_{s,t}(\psi)(y) = E^{y(0)}\big(\psi(Y^t) \mid Y^s = y\big)$, $s \leq t$. We prove the result for $t \leq \underline{t}$ by induction on $\underline{t}$. If $\underline{t} = 0$, then one has equality in the above. Assume the result for $t \leq \underline{t}$. Then

$$H_{\underline{t}+\tau}^N(\psi) = \frac{1}{N} \sum_{\alpha \sim \underline{t}} \psi\left(Y^\alpha_{\cdot \wedge (\underline{t}+\tau)}\right) N^\alpha,$$

and so

$$\begin{aligned}
E\left(H_{\underline{t}+\tau}^N(\psi)\right) &= E\left(\frac{1}{N} \sum_{\alpha \sim \underline{t}} \psi\left(Y^\alpha_{\cdot \wedge (\underline{t}+\tau)}\right) E\left(N^\alpha \mid \overline{\mathcal{F}}_{\underline{t}}\right)\right) \\
&= E\left(\frac{1}{N} \sum_{\alpha \sim \underline{t}} \psi\left(Y^\alpha_{\cdot \wedge (\underline{t}+\tau)}\right) \left(1 + g_N\left(Y^\alpha_{\underline{t}+\tau}\right)\tau\right)\right) \qquad \text{(by (II.3.5))} \\
&\leq (1 + \overline{g}\tau) E\left(\frac{1}{N} \sum_{\alpha \sim \underline{t}} P_{\underline{t},\underline{t}+\tau}\psi\left(Y^\alpha_{\cdot \wedge \underline{t}}\right)\right) \qquad \text{(by (II.3.3))} \\
&\leq e^{\overline{g}\tau} e^{\overline{g}\underline{t}} E^{X_0}\left(P_{\underline{t},\underline{t}+\tau}\psi(Y_{\cdot\wedge\underline{t}})\right) \qquad \text{(by induction hypothesis)} \\
&= e^{\overline{g}(\tau+\underline{t})} E^{X_0}\left(\psi(Y_{\cdot \wedge (\underline{t}+\tau)})\right).
\end{aligned}$$

Finally it should be clear from the above that the result follows for $t \in (\underline{t}, \underline{t} + \tau)$.
(b) Claim $e^{2g_\infty \underline{t}} X_{\underline{t}}^N(1)$ is an $\big(\overline{\mathcal{F}}_{\underline{t}}\big)$-submartingale for $N \geq N_0(g_\infty)$. From the above calculation we have

$$\begin{aligned}
&e^{2g_\infty(\underline{t}+\tau)} E\left(X_{\underline{t}+\tau}^N(1) \mid \overline{\mathcal{F}}_{\underline{t}}\right) \\
&\quad \geq e^{2g_\infty(\underline{t}+\tau)}(1 - g_\infty\tau) X_{\underline{t}}^N(1) \geq e^{2g_\infty \underline{t}} X_{\underline{t}}^N(1),
\end{aligned}$$

for $N \geq N_0(g_\infty)$. The weak L^1 inequality for non-negative submartingales and (a) now complete the proof. ∎

Remark II.3.4. It is clear from the above argument that if $g_N \equiv 0$, then equality holds in (a).

4. Tightness

We first specialize Theorem 3.1 of Jakubowski (1986), which gives necessary and sufficient conditions for tightness in $D(\mathbb{R}_+, S)$, to the case $S = M_F(E)$. As E is Polish, $M_F(E)$ is also Polish–see Theorem 3.1.7 of Ethier and Kurtz (1986) for the corresponding result for $M_1(E)$ from which the result follows easily. (An explicit complete metric is defined prior to Lemma II.7.5.) Therefore $D(M_F(E))$ is also Polish and Prohorov's theorem implies that a collection of laws on $D(M_F(E))$ is tight iff it is relatively compact.

Definition. A collection of processes $\{X^\alpha : \alpha \in I\}$ with paths in $D(S)$ is C-relatively compact in $D(S)$ iff it is relatively compact in $D(S)$ and all weak limit points are a.s. continuous.

Definition. $D_0 \subset C_b(E)$ is separating iff for any $\mu, \nu \in M_F(E)$, $\mu(\phi) = \nu(\phi)$ $\forall \phi \in D_0$ implies $\mu = \nu$.

Theorem II.4.1. Let D_0 be a separating class in $C_b(E)$ containing 1. A sequence of cadlag $M_F(E)$-valued processes $\{X^N, N \in \mathbb{N}\}$ is C-relatively compact in $D\big(M_F(E)\big)$ iff the following conditions hold:

(i) $\forall \varepsilon, T > 0$ there is a compact set $K_{T,\varepsilon}$ in E such that

$$\sup_N P\left(\sup_{t\le T} X_t^N(K_{T,\varepsilon}^c) > \varepsilon\right) < \varepsilon.$$

(ii) $\forall \phi \in D_0$, $\{X_\cdot^N(\phi) : N \in \mathbb{N}\}$ is C-relatively compact in $D(\mathbb{R}_+, \mathbb{R})$.

If, in addition, D_0 is closed under addition, then the above equivalence holds when ordinary relatively compactness in D replaces C-relative compactness in both the hypothesis and conclusion.

Remark. A version of this result is already implicit in Kurtz (1975) (see the Remark after Theorem 4.20). A proof of the sufficiency of the above conditions in the C-relatively compact setting is given at the end of this Section. All the ideas of the proof may be found in Theorem 3.9.1 and Corollary 3.9.2 of Ethier and Kurtz (1986).

Although the C-relatively compact version is the result we will need, a few words are in order about the result in the general setting. (i) essentially reduces the result to the case when E is compact. In this case it is not hard to see there is a countable subset $D_0' \subset D_0$ closed under addition such that $\psi(\mu) = (\mu(f))_{f\in D_0'}$ is a homeomorphism from $M_F(E)$ onto its image in $\mathbb{R}^{D_0'}$. The same is true of the map $X_t \to (X_t(\phi))_{\phi\in D_0'}$ from $D\big(M_F(E)\big)$ onto its image in $D(\mathbb{R}_+, \mathbb{R}^{D_0'})$. To complete the proof we must show $D(\mathbb{R}_+, \mathbb{R}^{D_0'})$ and $D(\mathbb{R}_+, \mathbb{R})^{D_0'}$ are homeomorphic. This is the step which requires D_0' to be closed under addition. As any scholar of the J_1-topology knows, $X^n \to X$ and $Y^n \to Y$ in $D(\mathbb{R}_+, \mathbb{R})$ need not imply $(X^n, Y^n) \to (X, Y)$ in $D(\mathbb{R}_+, \mathbb{R}^2)$, but it does if in addition $X^n + Y^n \to X + Y$. See Jakubowski (1986) for the details.

Notation. $\mathcal{F}_t^X = \bigcap_{u>t} \sigma(X_s : s \le u)$ denotes the right-continuous filtration generated by a process X. Let $A^f\phi = (A + f)\phi$ for $\phi \in \mathcal{D}(A), f \in C_b(E)$.

We will use Theorem II.4.1 to prove the following tightness result. Recall that our standing hypotheses (II.2.1), (II.2.2) and (II.3.1) are in force.

Proposition II.4.2. $\{X^N\}$ is C-relatively compact in $D\big(M_F(E)\big)$. Each weak limit point, X, satisfies

$$\forall\phi\in\mathcal{D}(A),\ M_t(\phi)=X_t(\phi)-X_0(\phi)-\int_0^t X_s(A^g\phi)ds \text{ is a continuous}$$

$(MP)_{X_0}$

$$\left(\mathcal{F}_t^X\right)\text{-martingale such that } M_0(\phi)=0 \text{ and } \langle M(\phi)\rangle_t=\int_0^t X_s(\gamma\phi^2)ds.$$

Proof. We take $D_0=\mathcal{D}(A)$ in Theorem II.4.1, which is a separating class by Corollary II.2.3. We first check (ii) in II.4.1.

Let $\phi\in\mathcal{D}(A)$ and define

$$M_t^\alpha=\phi(Y_t^\alpha)-\phi(Y_{\underline{t}}^\alpha)-\int_{\underline{t}}^t A\phi(Y_s^\alpha)ds,\ t\in\left[\underline{t},\underline{t}+N^{-1}\right],\ \underline{t}=|\alpha|/N,\alpha\in I.$$

Note that if we group the population at $\underline{s}+N^{-1}$ according to their parents at $\underline{s}$ we get

$$X^N_{\underline{s}+N^{-1}}(\phi)=\frac{1}{N}\sum_{\alpha\sim\underline{s}}\phi\left(Y^\alpha_{\underline{s}+N^{-1}}\right)N^\alpha.$$

Therefore,

$(II.4.1)$

$$\begin{aligned}X^N_{\underline{s}+N^{-1}}(\phi)-X^N_{\underline{s}}(\phi)&=\frac{1}{N}\sum_{\alpha\sim\underline{s}}\left[\phi\left(Y^\alpha_{\underline{s}+N^{-1}}\right)N^\alpha-\phi\left(Y^\alpha_{\underline{s}}\right)\right]\\
&=\frac{1}{N}\sum_{\alpha\sim\underline{s}}\phi\left(Y^\alpha_{\underline{s}+N^{-1}}\right)\left[N^\alpha-\left(1+g_N\left(Y^\alpha_{\underline{s}+N^{-1}}\right)N^{-1}\right)\right]+N^{-2}\sum_{\alpha\sim\underline{s}}\phi g_N\left(Y^\alpha_{\underline{s}+N^{-1}}\right)\\
&\quad+\frac{1}{N}\sum_{\alpha\sim\underline{s}}M^\alpha_{\underline{s}+N^{-1}}+\int_{\underline{s}}^{\underline{s}+N^{-1}}\frac{1}{N}\sum_{\alpha\sim\underline{s}}A\phi\left(Y^\alpha_s\right)ds\end{aligned}$$

and

$$\begin{aligned}X_t^N(\phi)-X_{\underline{t}}^N(\phi)&=\frac{1}{N}\sum_{\alpha\sim\underline{t}}(\phi\left(Y_t^\alpha\right)-\phi(Y_{\underline{t}}^\alpha))\\
&=\frac{1}{N}\sum_{\alpha\sim\underline{t}}M_t^\alpha+\int_{\underline{t}}^t\frac{1}{N}\sum_{\alpha\sim\underline{t}}A\phi\left(Y_s^\alpha\right)ds.\end{aligned}\qquad (II.4.2)$$

Sum (II.4.1) over $\underline{s} < \underline{t}$ and then add (II.4.2) to arrive at

$$X_t^N(\phi) = X_0^N(\phi) + \frac{1}{N}\sum_{\underline{s}<\underline{t}}\sum_{\alpha\sim\underline{s}} \phi\left(Y^\alpha_{\underline{s}+N^{-1}}\right)\left(N^\alpha - \left(1 + g_N(Y^\alpha_{\underline{s}+N^{-1}})N^{-1}\right)\right)$$
$$+ \int_0^{\underline{t}} X_{\underline{s}}^N(\phi g_N)ds + N^{-2}\sum_{\underline{s}<\underline{t}}\sum_{\alpha\sim\underline{s}}\left(\phi g_N\left(Y^\alpha_{\underline{s}+N^{-1}}\right) - \phi g_N(Y^\alpha_{\underline{s}})\right)$$
$$+ \left[\frac{1}{N}\sum_{\underline{s}<\underline{t}}\sum_{\alpha\sim\underline{s}} M^\alpha_{\underline{s}+N^{-1}} + \frac{1}{N}\sum_{\alpha\sim\underline{t}} M_t^\alpha\right] + \int_0^t X_s^N(A\phi)ds,$$

and therefore

$$(MP)^N \qquad X_t^N(\phi) = X_0^N(\phi) + M_{\underline{t}}^{b,N}(\phi) + \int_0^{\underline{t}} X_{\underline{s}}^N(\phi g_N)ds + \delta_{\underline{t}}^N(\phi)$$
$$+ M_t^{s,N}(\phi) + \int_0^t X_s^N(A\phi)ds.$$

In the last line, the terms $M_{\underline{t}}^{b,N}(\phi)$, $\delta_{\underline{t}}^N(\phi)$ and $M_t^{s,N}(\phi)$ are defined to be the corresponding terms in the previous expression.

We start by handling the error term δ^N.

Lemma II.4.3. $\sup_{\underline{t}\le K} |\delta_{\underline{t}}^N(\phi)| \xrightarrow{L^1} 0$ as $N \to \infty \quad \forall K \in \mathbb{N}$.

Proof. If $h_N(y) \equiv E^y\big(|g_N\phi(Y_{N^{-1}}) - g_N\phi(Y_0)|\big)$ then

$$E\left(\sup_{\underline{t}\le K}|\delta_{\underline{t}}^N(\phi)|\right) \le \sum_{\underline{s}<K} N^{-1} E\left(\frac{1}{N}\sum_{\alpha\sim\underline{s}} E\left(\left|g_N\phi(Y^\alpha_{\underline{s}+N^{-1}}) - g_N\phi(Y^\alpha_{\underline{s}})\right| \mid \mathcal{F}_{\underline{s}}\right)\right)$$
$$= \int_0^K E\left(X_{\underline{s}}^N(h_N)\right)ds \qquad \text{(by (II.3.3))}$$
$$\le e^{g_\infty K}\int_0^K E^{X_0}\left(h_N(Y_{\underline{s}})\right)ds \qquad \text{(Lemma II.3.3(a))}$$
$$= e^{g_\infty K} E^{X_0}\Big(\int_0^K |g_N\phi(Y_{\underline{s}+N^{-1}}) - g_N\phi(Y_{\underline{s}})|\,ds\Big).$$

Now since $\{Y_s, Y_{s-} : s \le K\}$ is a.s. compact, (II.3.1)(a) and Dominated Convergence show that for some $\eta_N \to 0$,

$$\begin{aligned}
E\Big(\sup_{\underline{t}\le K} |\delta^N_{\underline{t}}(\phi)|\Big) &\le e^{g_\infty K} E^{X_0}\Big(\int_0^K |g\phi(Y_{\underline{s}+N^{-1}}) - g\phi(Y_{\underline{s}})|\, ds\Big) + \eta_N \\
&\to e^{g_\infty K} E^{X_0}\Big(\int_0^K |g\phi(Y_s) - g\phi(Y_{s-})|\, ds\Big) \qquad \text{as } N \to \infty \\
&= 0. \quad \blacksquare
\end{aligned}$$

Use (II.3.3) and argue as in the proof of Proposition II.2.1 to see that

$$(M^\alpha_t, \mathcal{F}_t)_{t\in[\underline{t},\underline{t}+N^{-1}]} \text{ is a martingale.}$$

This and the fact that $\{\alpha \sim \underline{t}\} \in \mathcal{F}_{\underline{t}}$ (recall Exercise II.3.2) easily imply that $(M^{s,N}_t(\phi), \mathcal{F}_t)_{t\ge 0}$ is a martingale. Lemma II.3.3 (a) with $\phi \equiv 1$ implies integrability. Perhaps somewhat surprisingly, we now show that $M^{s,N}(\phi)$ will not contribute to the limit as $N \to \infty$. This is essentially a Strong Law effect. A moment's thought shows that the fact that Y has no fixed time discontinuities (by (QLC)) must play an implicit role in the proof as the following result fails without it.

Lemma II.4.4. $\sup_{t\le K} |M^{s,N}_t(\phi)| \xrightarrow{L^2} 0$ as $N \to \infty$ $\forall K > 0$.

Proof. Let $h_N(y) = E^y\big([\phi(Y_{1/N}) - \phi(Y_0)]^2\big)$. The definition of M^α_t and an easy orthogonality argument shows that for $K \in \mathbb{N}$

$$\begin{aligned}
E\left(M^{s,N}_K(\phi)^2\right) &= N^{-2} \sum_{\underline{s}<K} E\left(\sum_{\alpha\sim\underline{s}} E\left((M^\alpha_{\underline{s}+N^{-1}})^2 \mid \mathcal{F}_{\underline{s}}\right)\right) \\
&\le 2N^{-2} \sum_{\underline{s}<K} \left[E\left(\sum_{\alpha\sim\underline{s}} \left(h_N(Y^\alpha_{\underline{s}}) + \|A\phi\|^2_\infty N^{-2}\right)\right)\right] \qquad \text{(by (II.3.3))} \\
&\le 2E\left(\int_0^K X^N_{\underline{s}}(h_N) + \|A\phi\|^2_\infty N^{-2} X^N_{\underline{s}}(1) ds\right) \\
&\le 2e^{g_\infty K}\left[E^{X_0}\left(\int_0^K \left(\phi(Y_{\underline{s}+N^{-1}}) - \phi(Y_{\underline{s}})\right)^2 ds\right) + KN^{-2}\|A\phi\|^2_\infty X_0(1)\right].
\end{aligned}$$

In the last line we have used Lemma II.3.3(a) and argued as in the proof of Lemma II.4.3. As in that result the above expression approaches 0 as $N \to \infty$ by Dominated Convergence. This proves the above expectation goes to 0 and the result follows by the strong L^2 inequality for martingales. $\blacksquare$

Recall that $T = T_N = \{j/N : j \in \mathbb{Z}^+\}$. We claim $(M^{b,N}_{\underline{t}}(\phi), \overline{\mathcal{F}}_{\underline{t}})_{\underline{t}\in T}$ is also a martingale. To see this note that

$$\begin{aligned} &E\left(M^{b,N}_{\underline{t}+N^{-1}}(\phi) - M^{b,N}_{\underline{t}}(\phi) \mid \overline{\mathcal{F}}_{\underline{t}}\right) \\ &= \frac{1}{N}\sum_{\alpha\sim\underline{t}} \phi\left(Y^{\alpha}_{\underline{t}+N^{-1}}\right) E\left(N^{\alpha} - \left(1 + g_N(Y^{\alpha}_{\underline{t}+N^{-1}})N^{-1}\right) \mid \overline{\mathcal{F}}_{\underline{t}}\right) \\ &= 0 \end{aligned}$$

by our definition of the conditional law of N^{α}. Integrability of this increment is again clear from Lemma II.3.3(a) and $E(|N^{\alpha}|) \le C$. In view of the fact that our spatial martingales vanish in the limit we expect that the martingale $M_t(\phi)$ in $(MP)_{X_0}$ must arise from the "branching martingales" $M^{b,N}(\phi)$.

To analyze $M^{b,N}$ we use the following "well-known" result.

Lemma II.4.5. Let $\left(M^N_{\underline{t}}, \overline{\mathcal{F}}^N_{\underline{t}}\right)_{\underline{t}\in T_N}$ be martingales with $M^N_0 = 0$. Let $\langle M^N\rangle_{\underline{t}} = \sum_{0\le\underline{s}<\underline{t}} E\big((M^N_{\underline{s}+N^{-1}} - M^N_{\underline{s}})^2 \mid \overline{\mathcal{F}}^N_{\underline{s}}\big)$, and extend $M^N_{\cdot}$ and $\langle M^N\rangle_{\cdot}$ to $\mathbb{R}_+$ as right-continuous step functions.

(a) If $\{\langle M^N\rangle_{\cdot} : N \in \mathbb{N}\}$ is C-relatively compact in $D(\mathbb{R})$ and

$$(II.4.3) \qquad \sup_{0\le\underline{t}\le K} |M^N(\underline{t}+N^{-1}) - M^N(\underline{t})| \xrightarrow{P} 0 \quad \text{as} \quad N\to\infty \quad \forall K > 0,$$

then $\{M^N\}$ is C-relatively compact in $D(\mathbb{R})$.

(b) If, in addition,

$$(II.4.4) \qquad \left\{\left(M^N_{\underline{t}}\right)^2 + \langle M^N\rangle_{\underline{t}} : N \in \mathbb{N}\right\} \quad \text{is uniformly integrable} \quad \forall \underline{t} \in T,$$

then $M^{N_k} \overset{w}{\Rightarrow} M$ in $D(\mathbb{R})$ implies M is a continuous L^2 martingale and $(M^{N_k}, \langle M^{N_k}\rangle) \overset{w}{\Rightarrow} (M, \langle M\rangle)$ in $D(\mathbb{R})^2$.

(c) Under (II.4.4), the converse to (a) holds.

Proof. (a) is immediate from Theorems VI.4.13 and VI.3.26 of Jacod-Shiryaev (1987). A nonstandard proof (and statement) of the entire result may also be found in Theorems 8.5 and 6.7 of Hoover-Perkins (1983). (b) remains valid without the $\langle M^N\rangle_{\underline{t}}$ term in (II.4.4) but the proof with this condition becomes completely elementary as the reader can easily check. ∎

The key ingredient in the above result is a predictable square function inequality of Burkholder (1973):

(PSF)

$\exists c = c(c_0)$ such that if $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, increasing, $\phi(0) = 0$ and $\phi(2\lambda) \le c_0\phi(\lambda)$ for all $\lambda \ge 0$, $(M_n, \mathcal{F}_n)$ is a martingale, $M^*_n = \sup_{k\le n}|M_k|$, $\langle M\rangle_n = \sum_{k=1}^{n} E\left((M_k - M_{k-1})^2 \mid \mathcal{F}_{k-1}\right) + E(M_0^2)$ and $d^*_n = \max_{1\le k\le n}|M_k - M_{k-1}|$, then $E\left(\phi(M^*_n)\right) \le c\left[E\left(\phi\left(\langle M\rangle_n^{1/2}\right) + \phi(d^*_n)\right)\right]$.

To apply the above Lemma to $M_{\underline{t}}^{b,N}(\phi)$, first note that by (II.3.2)(c), if $\alpha \sim \underline{s}$, $\beta \sim \underline{s}$, and $\alpha \neq \beta$, then N^α and N^β are conditionally independent given $\bar{\mathcal{F}}_{\underline{s}}$. The resulting orthogonality then shows that

$$\begin{aligned} &E\left(\left(M^{b,N}_{\underline{s}+N^{-1}}(\phi) - M^{b,N}_{\underline{s}}(\phi)\right)^2 \mid \overline{\mathcal{F}}_{\underline{s}}\right) \\ &\quad = \frac{1}{N^2}\sum_{\alpha\sim\underline{s}} \phi(Y^\alpha_{\underline{s}+N^{-1}})^2 E\left(\left(N^\alpha - (1+g_N(Y^\alpha_{\underline{s}+N^{-1}})N^{-1})\right)^2 \mid \overline{\mathcal{F}}_{\underline{s}}\right) \\ &\quad = \frac{1}{N^2}\sum_{\alpha\sim\underline{s}} \phi(Y^\alpha_{\underline{s}+N^{-1}})^2 \gamma_N(Y^\alpha_{\underline{s}+N^{-1}}). \end{aligned}$$

Sum over $\underline{s} < \underline{t}$ to see that

$$(II.4.5) \qquad \langle M^{b,N}(\phi)\rangle_{\underline{t}} = \int_0^{\underline{t}} X^N_{\underline{s}}(\phi^2\gamma_N)ds + \varepsilon^N_{\underline{t}}(\phi) \le c_\phi \int_0^t X^N_{\underline{s}}(1)ds,$$

where

$$\varepsilon^N_{\underline{t}}(\phi) = N^{-2}\sum_{\underline{s}<\underline{t}}\sum_{\alpha\sim\underline{s}} \phi^2\gamma_N(Y^\alpha_{\underline{s}+N^{-1}}) - \phi^2\gamma_N(Y^\alpha_{\underline{s}}).$$

Just as in Lemma II.4.3, for any $K > 0$,

$$(II.4.6) \qquad \sup_{\underline{t}\le K} |\varepsilon^N_{\underline{t}}(\phi)| \xrightarrow{L^1} 0 \quad \text{as} \quad N \to \infty.$$

We also see from the above that for $0 \le \underline{s} < \underline{t} \le K$,

$$\langle M^{b,N}(\phi)\rangle_{\underline{t}} - \langle M^{b,N}(\phi)\rangle_{\underline{s}} \le c \sup_{\underline{s}\le K} X^N_{\underline{s}}(1)|\underline{t} - \underline{s}|,$$

which in view of Lemma II.3.3(b) and Arzela-Ascoli implies the C-relative compactness of $\{\langle M^{b,N}(\phi)\rangle_{.} : N \in \mathbb{N}\}$ in $D(\mathbb{R})$. To verify (II.4.3) and (II.4.4) we will use the following result, whose proof we defer.

Lemma II.4.6. $\sup_N E\left(\sup_{\underline{t}\le K} X^N_t(1)^2\right) < \infty$ for any $K > 0$.

Exercise II.4.1 (a) If δ is as in (II.3.1(c)) use (PSF) to show

$$\lim_{N\to\infty} E\left(\sum_{\underline{t}\le K} |\Delta M^{b,N}(\phi)(\underline{t})|^{2+\delta}\right) = 0 \qquad \forall K > 0.$$

Hint: Conditional on $\overline{\mathcal{F}}_{\underline{t}}$, $\Delta M^{b,N}(\phi)(\underline{t})$ is a sum of mean 0 independent r.v.'s to which one may apply (PSF).

(b) Use (a), (PSF), and Lemma II.4.6 to show that $\forall K > 0$

$$\sup_N E\left(\sup_{\underline{t}\le K} |M^{b,N}_{\underline{t}}(\phi)|^{2+\delta}\right) < \infty$$

and, in particular, $\left\{\sup\limits_{\underline{t}\le K}|M^{b,N}_{\underline{t}}(\phi)|^2 : N\in\mathbb{N}\right\}$ is uniformly integrable.

The uniform integrability of $\left\{\langle M^{b,N}(\phi)\rangle_{\underline{t}} : N\in\mathbb{N}\right\}$ is clear from (II.4.5) and Lemma II.4.6, and so the above Exercise allows us to apply the relative compactness lemma for martingales (Lemma II.4.5) to conclude:
(M^b)

$\{M^{b,N}_{\cdot}(\phi) : N\in\mathbb{N}\}$ are C-relatively compact in D, and all limit points are continuous L^2 martingales. If $M^{b,N_k}_{\cdot}(\phi)\overset{w}{\Rightarrow}M_{\cdot}(\phi)$, then $\left(M^{b,N_k}_{\cdot}(\phi),\langle M^{b,N_k}(\phi)\rangle_{\cdot}\right)\overset{w}{\Rightarrow}(M_{\cdot}(\phi),\langle M(\phi)\rangle)$ in D^2.

Proof of Lemma II.4.6. Let $\phi = 1$ in $(MP)^N$ and combine the terms $\int\limits_0^{\underline{t}} X^N_{\underline{s}}(g_N)ds+\delta^N_{\underline{t}}(1)$ to see that (recall that $g_\infty=\sup_N\|g_N\|$)

$$X^N_{\underline{t}}(1)\le X^N_0(1)+M^{b,N}_{\underline{t}}(1)+g_\infty\int\limits_0^{\underline{t}} X^N_{\underline{s}}(1)ds. \tag{II.4.7}$$

Doob's strong L^2 inequality and Lemma II.3.3(a) imply

$$\begin{aligned}E\left(\sup_{\underline{t}\le K}M^{b,N}_{\underline{t}}(1)^2\right)&\le cE\left(\langle M^{b,N}(1)\rangle_K\right)\\ &\le cE\left(\int\limits_0^K X^N_{\underline{s}}(1)ds\right)\qquad\text{(by (II.4.5))}\\ &\le c(K)X_0(1).\end{aligned}\tag{II.4.8}$$

Jensen's inequality and (II.4.7) show that for $u\le K$

$$\sup_{t\le u}X^N_t(1)^2\le cX^N_0(1)^2+c\sup_{\underline{t}\le K}M^{b,N}_{\underline{t}}(1)^2+cg_\infty^2K\int\limits_0^{\underline{u}}X^N_{\underline{s}}(1)^2ds.$$

Consider the first $\underline{u}\le K$ at which the mean of the left side is infinite. If such a $\underline{u}$ exists, the last integral has finite mean and so does the right-hand side. This contradiction shows $E\big(\sup\limits_{t\le K}X^N_t(1)^2\big)<\infty$ and a simple Gronwall argument now gives a bound uniform in N, namely for all $u\le K$,

$$E\left(\sup_{t\le u}X^N_t(1)^2\right)\le c\left[\sup_N E\left(X^N_0(1)^2\right)+c(K)X_0(1)\right]e^{cg_\infty^2Ku}.\quad\blacksquare$$

To complete the verification of (ii) in Theorem II.4.1, return to $(MP)^N$. Recall from Remark II.3.2 that $X^N_0(\phi)\overset{P}{\to}X_0(\phi)$ as $N\to\infty$. The Arzela-Ascoli Theorem and Lemma II.3.3(b) show that $\int\limits_0^{\underline{t}}X^N_{\underline{s}}(\phi g_N)ds$ and $\int\limits_0^t X^N_s(A\phi)ds$ are C-relatively compact sequences in $D(\mathbb{R})$. Therefore $(MP)^N$, (M^b) and Lemmas II.4.3 and II.4.4 show that $\{X_\cdot^N(\phi)\}$ is C-relatively compact in $D(\mathbb{R})$ for each ϕ in $\mathcal{D}(A)$, and (ii) is verified.

We now give the

Proof of the Compact Containment Condition (i) in Theorem II.4.1. Let $\varepsilon, T > 0$ and $\eta = \eta(\varepsilon, T) > 0$ (η will be chosen below). As any probability on a Polish space is tight, we may choose a compact set $K_0 \subset D(E)$ so that $P^{X_0}(Y \in K_0^c) < \eta$. Let $K = \{y_t, y_{t-} : t \le T, y \in K_0\}$. It is easy to see K is compact in E (note that if $t_n \to t$ and $y_n \to y$, then $y_{n_k}(t_{n_k}) \to y(t)$ or $y(t-)$ for some subsequence $\{n_k\}$ and similarly for $y_{n_k}(t_{n_k}-)$). Clearly

$$P^{X_0}\left(Y_t \in K^c \quad \text{or} \quad Y_{t-} \in K^c \quad \exists t \le T\right) < \eta.$$

Let

$$\begin{aligned} R_t^N &= e^{2g_\infty \underline{t}} H_t^N\{y : y(s) \in K^c \quad \text{for some} \quad s \le t\}. \\ &= e^{2g_\infty \underline{t}} \frac{1}{N} \sum_{\alpha \sim t} \sup_{s \le t} 1_{K^c}(Y_s^\alpha). \end{aligned}$$

Claim R_t^N is an $\mathcal{F}_t^N$-submartingale for $N \ge N_0$. As $R_\cdot^N$ is increasing on $[\underline{t}, \underline{t} + \tau)$, it suffices to show that

$$E\left(R_{\underline{t}}^N - R_{\underline{t}-}^N \mid \mathcal{F}_{\underline{t}-}^N\right) \ge 0 \quad \text{a.s.} \tag{II.4.9}$$

We have

$$\begin{aligned} R_{\underline{t}}^N - R_{\underline{t}-}^N &= \frac{1}{N} \sum_{\alpha \sim \underline{t} - N^{-1}} \left[e^{2g_\infty \underline{t}} \sup_{s \le \underline{t}} 1_{K^c}(Y_s^\alpha) N^\alpha - e^{2g_\infty(\underline{t} - N^{-1})} \sup_{s < \underline{t}} 1_{K^c}(Y_s^\alpha) \right] \\ &\ge \frac{1}{N} \sum_{\alpha \sim \underline{t} - N^{-1}} \left[e^{2g_\infty \underline{t}} N^\alpha - e^{2g_\infty(\underline{t} - N^{-1})} \right] \sup_{s < \underline{t}} 1_{K^c}(Y_s^\alpha) \\ &= \frac{1}{N} \sum_{\alpha \sim \underline{t} - N^{-1}} e^{2g_\infty \underline{t}} \left[N^\alpha - g_N(Y_{\underline{t}}^\alpha)/N - 1 \right] \sup_{s < \underline{t}} 1_{K^c}(Y_s^\alpha) \\ &\quad + \frac{1}{N} \sum_{\alpha \sim \underline{t} - N^{-1}} e^{2g_\infty \underline{t}} \left[g_N(Y_{\underline{t}}^\alpha)/N + 1 - e^{-2g_\infty/N} \right] \sup_{s < \underline{t}} 1_{K^c}(Y_s^\alpha). \end{aligned}$$

The conditional expectation of the first term with respect to $\mathcal{F}_{\underline{t}-}^N$ is 0. The second term is at least

$$\frac{1}{N} \sum_{\alpha \sim \underline{t} - N^{-1}} e^{2g_\infty \underline{t}} \left[-g_\infty/N + 1 - e^{-2g_\infty/N} \right] \sup_{s < \underline{t}} 1_{K^c}(Y_s^\alpha) \ge 0, \quad \text{for} \quad N \ge N_0(g_\infty),$$

and (II.4.9) is proved. Now use the weak L^1 inequality for submartingales and

Lemma II.3.3 (a) to see that for $N \ge N_0(g_\infty)$

$$\begin{aligned} P\left(\sup_{t\le T} X_t^N(K^c) > \varepsilon\right) &\le P\left(\sup_{t\le T} R_t^N > \varepsilon\right) \\ &\le \varepsilon^{-1} E(R_T^N) \\ &\le e^{2g_\infty T}\varepsilon^{-1} e^{\overline{g}T} P^{X_0}\left(Y_s \in K^c \quad \exists s \le T\right) \\ &< \varepsilon \end{aligned}$$

by an appropriate choice of $\eta = \eta(\varepsilon, T)$. It is trivial to enlarge K if necessary to accommodate $N < N_0(g_\infty)$, e.g. by using the converse of Theorem II.4.1 and the fact that each P_N is trivially tight. ∎

Completion of Proof of Proposition II.4.2. By Theorem II.4.1 $\{X_\cdot^N\}$ is C-relatively compact in $D\big(M_F(E)\big)$. To complete the proof of Proposition II.4.2 we must verify that all limit points satisfy $(MP)_{X_0}$. Assume $X^{N_k} \overset{w}{\Rightarrow} X$. By a theorem of Skorohod (see Theorem 3.1.8 of Ethier-Kurtz (1986)) we may assume X^{N_k}, X are defined on a common probability space and $X^{N_k} \to X$ in $D\big(M_F(E)\big)$ a.s. Let $\phi \in \mathcal{D}(A)$. Note that $X \to \int_0^t X_r(A\phi)dr$ and $X \to \int_0^t X_r(g\phi)dr$ are continuous maps from $D\big(\mathbb{R}_+, M_F(E)\big)$ to $C(\mathbb{R}_+, \mathbb{R})$, and $X \to X_\cdot(\phi)$ is a continuous function from $D(M_F(E))$ to $D(\mathbb{R})$. This, Lemmas II.4.3 and II.4.4, the convergence in probability of $X_0^{N_k}(\phi)$ to $X_0(\phi)$ (by (II.3.6)), the uniform convergence on compacts of g_{N_k} to g and condition (i) in Theorem II.4.1 allow us to take limits in $(MP)^{N_k}$ and conclude that $M_\cdot^{b,N_k}(\phi) \to M_\cdot(\phi)$ in $D(\mathbb{R}_+, \mathbb{R})$ in probability where

$$X_t(\phi) = X_0(\phi) + M_t(\phi) + \int_0^t X_s(A\phi + g\phi)ds.$$

By (M^b), $M_\cdot(\phi)$ is a continuous square integrable martingale. In addition, (M^b) together with (II.4.5), (II.4.6), $\gamma_{N_k} \overset{ucb}{\to} \gamma$ and the compact containment condition (i) allow one to conclude

$$\langle M(\phi)\rangle_t = P - \lim_{k\to\infty} \int_0^{\underline{t}} X_{\underline{s}}^{N_k}(\phi^2\gamma_{N_k})ds = \int_0^t X_s(\phi^2\gamma)ds.$$

To complete the derivation of $(MP)_{X_0}$ we must show that $M_t(\phi)$ is an $(\mathcal{F}_t^X)$-martingale and not just a martingale with respect to its own filtration. To see this let $s_1 \le s_2 \ldots \le s_n \le s < t$, $\psi : M_F(E)^n \to \mathbb{R}$ be bounded and continuous and use Exercise II.4.1 (b) to take limits in

$$E\left(\left(M_{\underline{t}}^{b,N_k}(\phi) - M_{\underline{s}}^{b,N_k}(\phi)\right)\psi\left(X_{s_1}^{N_k}, \ldots, X_{s_n}^{N_k}\right)\right) = 0.$$

This shows that X satisfies $(MP)_{X_0}$ and thus completes the proof of Proposition II.4.2. ∎

Proof of Theorem II.4.1. We only prove the sufficiency of the two conditions as the necessity is quite easy. Let d be a complete metric on $M_F(E)$ and if $x \in D(M_F(E))$ and $\delta, T > 0$, set

$$w(x, \delta, T) = \sup\{d(x(t), x(s)) : \ s, t \le T, \ |s - t| \le \delta\}.$$

A standard result for general Polish state spaces states that $\{X^N : N \in \mathbb{N}\}$ is C-relatively compact if and only if

$$(II.4.10) \qquad \forall \varepsilon, T > 0 \text{ there is a compact set } K^0_{\varepsilon,T} \subset M_F(E) \text{ such that } \sup_N P(X^N(t) \notin K^0_{\varepsilon,T} \text{ for some } t \le T) \le \varepsilon,$$

and

$$(II.4.11) \ \forall \varepsilon, T > 0 \text{ there is a } \delta > 0 \text{ so that } \limsup_{N \to \infty} P(w(X^N, \delta, T) \ge \varepsilon) \le \varepsilon.$$

For example, this follows from Corollary 3.7.4, Remark 3.7.3, and Theorem 3.10.2 of Ethier and Kurtz (1986).

We first verify the compact containment condition (II.4.10). Let $\varepsilon, T > 0$. By condition (i) there are compact subsets K_m of E so that

$$\sup_N P(\sup_{t \le T} X^N_t(K^c_m) > 2^{-m}) < \varepsilon 2^{-m-1}.$$

Take $\phi = 1$ in condition (ii) and use (II.4.10) for the real-valued processes $X^N_\cdot(1)$ to see there is an $R = R(\varepsilon, T)$ so that

$$\sup_N P(\sup_{t \le T} X^N_t(1) > R) < \varepsilon/2.$$

Define

$$C^0 = \{\mu \in M_F(E) : \mu(K^c_m) \le 2^{-m} \text{ for all } m \in \mathbb{N}, \text{ and } \mu(1) \le R\}.$$

Then the choice of R and K_m imply that

$$P(X^N_t \notin C^0 \text{ for some } t \le T) < \varepsilon.$$

To verify compactness of $\overline{C^0}$, let $\{\mu_n\}$ be a sequence in C^0. To find a weakly convergent subsequence we may assume that $\inf \mu_n(E) > \delta > 0$. Tightness of $\{\mu_n/\mu_n(E)\}$ is now clear, and so by Prohorov's theorem there is a subsequence $\{n_k\}$over which these normalized measures converge weakly. As the total masses are bounded by R, we may take a further subsequence to ensure convergence of the total masses and hence obtain weak convergence of $\{\mu_{n_k}\}$. It follows that $K_0 = \overline{C^0}$ is compact and so will satisfy (II.4.10).

The next step is to show

$$(II.4.12) \ \ \forall f \in C_b(M_F(E)), \ \{f \circ X^N : N \in \mathbb{N}\} \text{ is } C\text{-relatively compact in } D(\mathbb{R}).$$

Let $f \in C_b(M_F(E))$ and $\varepsilon, T > 0$. Choose K^0 as in (II.4.10) and define

$$A = \Big\{h : M_F(E) \to \mathbb{R} \ : \ h(\mu) = \sum_{i=1}^{k} a_i \prod_{j=1}^{m_i} \mu(f_{i,j}),\ a_i \in \mathbb{R},\ f_{i,j} \in D_0, k, m_i \in \mathbb{Z}_+\Big\}$$
$$\subset C_b(M_F(E)).$$

Then A is an algebra containing the constant functions and separating points in $M_F(E)$. By Stone-Weierstrass there is an $h \in A$ so that $\sup_{\mu \in K^0} |h(\mu) - f(\mu)| < \varepsilon$. If $\{Y^N\}$ and $\{Z^N\}$ are C-relatively compact in $D(\mathbb{R})$ then so are $\{aY^N + bZ^N\}$ and $\{Y^N Z^N\}$ for any $a, b \in \mathbb{R}$. This is easy to show using (II.4.10) and (II.4.11), for example (but is false for ordinary relative compactness in $D(\mathbb{R})$). Therefore condition (ii) of the Theorem implies that $\{h \circ X^N\}$ is C-relatively compact and by (II.4.11) there is a $\delta > 0$ so that

$$(II.4.13) \qquad \limsup_{N\to\infty} P(w(h \circ X^N, \delta, T) \geq \varepsilon) \leq \varepsilon.$$

If $s, t \leq T$ and $|t - s| \leq \delta$, then

$$|f(X_t^N) - f(X_s^N)| \leq 2\|f\|_\infty 1(X^N([0,T]) \not\subset K^0) + 2 \sup_{\mu \in K^0} |h(\mu) - f(\mu)| + |h(X_t^N) - h(X_s^N)|,$$

and so,

$$w(f \circ X^N, \delta, T) \leq 2\|f\|_\infty 1(X^N([0,T]) \not\subset K^0) + 2\varepsilon + w(h \circ X^N, \delta, T).$$

Therefore

$$\begin{aligned}
&\limsup_{N\to\infty} P(w(f \circ X^N, \delta, T) \geq 3\varepsilon) \\
&\quad \leq \limsup_{N\to\infty} P(X^N([0,T]) \not\subset K^0) + \limsup_{N\to\infty} P(w(h \circ X^N, \delta, T) \geq \varepsilon) \leq 2\varepsilon,
\end{aligned}$$

the last by (II.4.13) and the choice of K^0. We have verified (II.4.11) with $\{f \circ X^N\}$ in place of $\{X^N\}$, and as (II.4.10) is trivial for this process, (II.4.12) follows.

It remains to verify (II.4.11). We may assume d is bounded by 1. Let $\varepsilon, T > 0$, and K^0 is as in (II.4.10). Choose $\mu_i \in K^0$, $i \leq M$, so that $K^0 \subset \cup_{i=1}^M B(\mu_i, \varepsilon)$, and let $f_i(\mu) = d(\mu_i, \mu)$. Clearly $f_i \in C_b(M_F(E))$. We showed in the previous paragraph that there is a $\delta > 0$ so that

$$(II.4.14) \qquad \sum_{i=1}^{M} \limsup_{N\to\infty} P(w(f_i \circ X^N, \delta, T) \geq \varepsilon) \leq \varepsilon.$$

If $\mu, \nu \in K^0$, choose μ_j so that $d(\nu, \mu_j) < \varepsilon$. Then

$$\begin{aligned}
d(\mu, \nu) &\leq d(\mu, \mu_j) + d(\mu_j, \nu) \\
&\leq |d(\mu, \mu_j) - d(\mu_j, \nu)| + 2d(\mu_j, \nu) \\
&\leq \max_i |f_i(\mu) - f_i(\nu)| + 2\varepsilon.
\end{aligned}$$

Let $s, t \le T$, $|s-t| \le \delta$. Then the above inequality implies that

$$d(X_t^N, X_s^N) \le \max_i |f_i \circ X^N(t) - f_i \circ X^N(s)| + 2\varepsilon + 1(X^N([0,T]) \not\subset K^0),$$

and therefore,

$$w(X^N, \delta, T) \le \max_i w(f_i \circ X^N, \delta, T) + 2\varepsilon + 1(X^N([0,T]) \not\subset K^0).$$

It follows that

$$\begin{aligned}&\limsup_{N\to\infty} P(w(X^N, \delta, T) \ge 3\varepsilon)\\ &\qquad \le \limsup_{N\to\infty} P(\max_i w(f_i \circ X^N, \delta, T) \ge \varepsilon) + \limsup_{N\to\infty} P(X^N([0,T]) \not\subset K^0) \le 2\varepsilon,\end{aligned}$$

the last by (II.4.14) and the choice of K^0. This gives (II.4.11) and so the proof is complete. ∎

5. The Martingale Problem

In order to prove convergence in Proposition II.4.2 it suffices to show solutions to $(MP)_{X_0}$ are unique in law. We will show this is the case and state the main result of this section in Theorem II.5.1 below. Let g, $\gamma \in C_b(E)$ with $\gamma \ge 0$ as before. A is the weak generator of our BSMP, Y, satisfying (II.2.1) and (II.2.2). Recall that $A^g\phi = A\phi + g\phi$. $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ will denote a filtered probability space with $(\mathcal{F}_t)$ right-continuous.

Definition. Let ν be a probability on $M_F(E)$. An adapted a.s. continuous $M_F(E)$-valued process, X, on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ satisfies $(LMP)_\nu$ (or $(LMP)_\nu^{g,\gamma,A}$) iff

$$(LMP)_\nu \quad \begin{aligned}&X_0 \text{ has law } \nu \text{ and } \forall \phi \in \mathcal{D}(A)\ M_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(A^g\phi)ds\\ &\text{is an } (\mathcal{F}_t) - \text{local martingale such that } \langle M(\phi)\rangle_t = \int_0^t X_s(\gamma\phi^2)ds.\end{aligned}$$

Remark. If $\int X_0(1)d\nu(X_0) = \infty$, then the integrability of $M_t(1)$ may fail and so we need to work with a local martingale problem. We let $(MP)_\nu$ denote the corresponding martingale problem (i.e. $M_t(\phi)$ is an $(\mathcal{F}_t)$-martingale), thus slightly abusing the notation in Proposition II.4.2. That result shows that if $X_0 \in M_F(E)$, then any limit point of $\{X^N\}$ satisfies $(MP)_{\delta_{X_0}}$ on the canonical space of measure-valued paths.

Definition. $(LMP)_\nu$ is well-posed if a solution exists on some $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and the law of any solution (on $C(\mathbb{R}_+, M_F(E))$) is unique.

Notation. $\Omega_X = C(\mathbb{R}_+, M_F(E))$, $\mathcal{F}_X =$ Borel sets on Ω_X, $\Omega_D = D(\mathbb{R}_+, M_F(E))$.

Theorem II.5.1. (a) $(LMP)_\nu$ is well-posed $\forall \nu \in M_1(M_F(E))$.

(b) There is a family of probabilities $\{\mathbb{P}_{X_0} : X_0 \in M_F(E)\}$ on $(\Omega_X, \mathcal{F}_X)$ such that if $X_t(\omega) = \omega_t$, then

(i) $\mathbb{P}_\nu(\cdot) = \int \mathbb{P}_{X_0}(\cdot) d\nu(X_0)$ is the law of any solution to $(LMP)_\nu$ for any probability ν on $M_F(E)$.

(ii) $(\Omega_X, \mathcal{F}_X, \mathcal{F}_t^X, X, \mathbb{P}_{X_0})$ is a BSMP.

(iii) If $(Z_t)_{t\geq 0}$ satisfies $(LMP)_\nu$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and T is an a.s. finite $(\mathcal{F}_t)$-stopping time, then $\mathbb{P}\big(Z_{T+\cdot} \in A \mid \mathcal{F}_T\big)(\omega) = \mathbb{P}_{Z_T(\omega)}(A)$ a.s. $\forall A \in \mathcal{F}_X$.

(c) If (II.3.1) holds and $\{X^N\}$ are as in Proposition II.4.2, then

$$\mathbb{P}(X^N \in \cdot) \overset{w}{\Rightarrow} \mathbb{P}_{X_0} \quad \text{on} \quad \Omega_D.$$

(d) If $T_t F(X_0) = \mathbb{P}_{X_0}\big(F(X_t)\big)$, then $T_t : C_b\big(M_F(E)\big) \to C_b\big(M_F(E)\big)$.

The key step in the above is the uniqueness of solution to $(LMP)_\nu$. The remaining properties will be standard consequences of this and the method (duality) used to establish uniqueness (see, e.g., Theorem II.5.6 below). A process satisfying $(LMP)_\nu$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is called an $(\mathcal{F}_t)$-(A, γ, g)-Dawson-Watanabe superprocess, or $(\mathcal{F}_t)$-(Y, γ, g)-DW superprocess, with initial law ν, or, if $\nu = \delta_{X_0}$, starting at X_0.

The following standard monotone class lemma will be useful.

Lemma II.5.2. Let $\mathcal{H} \subset b\mathcal{F}$ be a linear class containing 1 and closed under $\overset{bp}{\to}$. Let $\mathcal{H}_0 \subset \mathcal{H}$ be closed under products. Then $\mathcal{H}$ contains all bounded $\sigma(\mathcal{H}_0)$-measurable functions.

Proof. See p. 497 of Ethier-Kurtz (1986). ∎

Let X satisfy $(LMP)_\nu$ on some $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Let $\mathcal{M}_{\text{loc}}$ be the space of continuous $(\mathcal{F}_t)$-local martingales such that $M_0 = 0$. Here processes which agree off an evanescent set are identified. Let $\mathcal{P}$ be the σ-field of $(\mathcal{F}_t)$-predictable sets in $\mathbb{R}_+ \times \Omega$, and define

$$\mathcal{L}^2_{\text{loc}} = \Big\{\psi : \mathbb{R}_+ \times \Omega \times E \to \mathbb{R} : \psi \text{ is } \mathcal{P} \times \mathcal{E}\text{−measurable}, \int_0^t X_s(\psi_s^2 \gamma) ds < \infty \quad \forall t > 0\Big\},$$

and

$$\mathcal{L}^2 = \{\psi \in \mathcal{L}^2_{\text{loc}} : E\Big(\int_0^t X_s(\psi_s^2\gamma) ds\Big) < \infty \ \forall t > 0\}.$$

A $\mathcal{P} \times \mathcal{E}$-measurable function ψ is simple (write $\psi \in \mathcal{S}$) iff

$$\psi(t, \omega, x) = \sum_{i=0}^{K-1} \psi_i(\omega)\phi_i(x) 1_{(t_i, t_{i+1}]}(t)$$

for some $\phi_i \in \mathcal{D}(A)$, $\psi_i \in b\mathcal{F}_{t_i}$ and $0 = t_0 < t_1 \ldots < t_K \leq \infty$. For such a ψ define

$$M_t(\psi) \equiv \int_0^t \int \psi(s, x) dM(s, x) = \sum_{i=0}^{K-1} \psi_i \left(M_{t\wedge t_{i+1}}(\phi_i) - M_{t\wedge t_i}(\phi_i)\right).$$

Then a standard argument shows that $M_t(\psi)$ is well-defined (i.e., independent of the choice of representation for ψ) and so $\psi \mapsto M(\psi)$ is clearly linear. If

$$\tilde{\psi}_i(s,\omega) = \psi_i(\omega)1_{(t_i,t_{i+1}]}(s),$$

then $\tilde{\psi}_i$ is $\mathcal{P}$-measurable and $M_t(\psi) = \sum_{i=0}^{K-1} \int_0^t \tilde{\psi}_i dM_s(\phi_i)$. Therefore $M_t(\psi)$ is in $\mathcal{M}_{\text{loc}}$ and a simple calculation gives

$$\langle M(\psi)\rangle_t = \int_0^t X_s\left(\gamma\psi_s^2\right) ds.$$

Lemma II.5.3. For any $\psi \in \mathcal{L}^2_{\text{loc}}$ there is a sequence $\{\psi_n\}$ in $\mathcal{S}$ such that

$$\mathbb{P}\left(\int_0^n \int (\psi_n - \psi)^2(s,\omega,x)\gamma(x)X_s(dx)ds > 2^{-n}\right) < 2^{-n}.$$

Proof. Let $\overline{\mathcal{S}}$ denote the set of bounded $\mathcal{P}\times\mathcal{E}$-measurable functions which can be approximated as above. $\overline{\mathcal{S}}$ is clearly closed under $\xrightarrow{bp}$. Since $\overline{D(A)}^{bp} = b\mathcal{E}$, $\overline{\mathcal{S}}$ contains

$$\psi(t,\omega,x) = \sum_{i=0}^{K-1} \psi_i(\omega,x)1_{(t_i,t_{i+1}]}(t)$$

where $0 = t_0 < \ldots < t_K \le \infty$ and $\psi_i(\omega,x) = f_i(\omega)\phi_i(x)$, $\phi_i \in b\mathcal{E}$, $f_i \in b\mathcal{F}_{t_i}$. Now apply Lemma II.5.2 to the class $\mathcal{H}$ of $\psi_i \in b(\mathcal{F}_{t_i}\times\mathcal{E})$ for which ψ as above is in $\overline{\mathcal{S}}$. Using $\mathcal{H}_0 = \{f_i(\omega)\phi_i(x) : f_i \in b\mathcal{F}_{t_i}, \phi_i \in b\mathcal{E}\}$, we see that $\psi(t,\omega,x) = \sum\limits_{i=0}^{K-1} \psi_i(\omega,x)1_{(t_i,t_{i+1}]}(t)$ is in $\overline{\mathcal{S}}$ for any $\psi_i \in b(\mathcal{F}_{t_i}\times\mathcal{E})$.

If $\psi \in b(\mathcal{P}\times\mathcal{E})$, then

$$\psi_n(s,\omega,x) = 2^n \int_{(i-1)2^{-n}}^{i2^{-n}} \psi(r,\omega,x)dr \quad \text{if} \quad s \in \left(i2^{-n},(i+1)2^{-n}\right] \qquad i = 1,2,\ldots$$

satisfies $\psi_n \in \overline{\mathcal{S}}$ by the above. For each (ω,x), $\psi_n(s,\omega,x) \to \psi(s,\omega,x)$ for Lebesgue a.a. s by Lebesgue's differentiation theorem (e.g. Theorem 8.8 of Rudin (1974)) and it follows easily that $\psi \in \overline{\mathcal{S}}$. Finally if $\psi \in \mathcal{L}^2_{\text{loc}}$, the obvious truncation argument and Dominated Convergence (set $\psi_n = (\psi \wedge n) \vee (-n)$) completes the proof. ∎

Proposition II.5.4. There is a unique linear extension of $M : \mathcal{S} \to \mathcal{M}_{\text{loc}}$ to a map $M : \mathcal{L}^2_{\text{loc}} \to \mathcal{M}_{\text{loc}}$ such that $\langle M(\psi)\rangle_t = \int\limits_0^t X_s(\gamma\psi_s^2)ds$ $\forall t \ge 0$ a.s. $\forall \psi \in \mathcal{L}^2_{\text{loc}}$. If $\psi \in \mathcal{L}^2$, then $M(\psi)$ is a square integrable $\mathcal{F}_t$-martingale.

Proof. Assume M satisfies the above properties and $\psi \in \mathcal{L}^2_{\text{loc}}$. Choose $\psi_n \in \mathcal{S}$ as in Lemma II.5.3. By linearity,

$$\langle M(\psi) - M(\psi_n)\rangle_n = \langle M(\psi - \psi_n)\rangle_n = \int_0^n X_s\Big(\gamma\big(\psi(s) - \psi_n(s)\big)^2\Big)\,ds$$
$$\le 2^{-n} \text{ w.p. } > 1 - 2^{-n}.$$

A standard square function inequality and the Borel-Cantelli Lemma imply

$$\sup_{t\le n}\big|M_t(\psi) - M_t(\psi_n)\big| \to 0 \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

This proves uniqueness and shows how we must define the extension. The details showing that this extension has the required properties are standard and left for the reader. Finally it is easy to use the dominated convergence theorem to see that if $M \in \mathcal{M}_{\text{loc}}$ satisfies $E(\langle M\rangle_t) < \infty$ for all $t > 0$ then M is a square integrable martingale. This proves the final assertion for $\psi \in \mathcal{L}^2$. ∎

Remarks. II.5.5. (1) By polarization if ϕ, $\psi \in \mathcal{L}^2_{\text{loc}}$,

$$\langle M(\phi), M(\psi)\rangle_t = \int_0^t X_s(\gamma\phi_s\psi_s)ds.$$

In particular if A_1 and A_2 are disjoint sets in $\mathcal{E}$ and $M(A_i) = M(1_{A_i})$, then $\langle M(A_1), M(A_2)\rangle_t = 0$ and so M_t is an orthogonal (local) martingale measure in the sense of Ch. 2 of Walsh (1986) where the reader can find more general constructions of this type.

(2) If $\int X_0(1)d\nu(X_0) < \infty$, then take $\phi \equiv 1$ in $(LMP)_\nu$ and use Gronwall's and Fatou's Lemmas to see that $E\big(X_t(1)\big) \le E\big(X_0(1)\big)e^{\tilde{g}t}$ where $\tilde{g} = \sup_x g(x)$. If ψ is $\mathcal{P} \times \mathcal{E}$ measurable and bounded on $[0, T] \times \Omega \times E$ for all $T > 0$, then the above shows that $\psi \in \mathcal{L}^2$ and so $M_t(\psi)$ is an L^2 martingale.

Let $\mathcal{M}_F$ denote the Borel σ-field on $M_F(E)$.

Theorem II.5.6. Assume:

(H_1) $\forall X_0 \in M_F(E)$ there is a solution to $(LMP)_{\delta_{X_0}}$.

(H_2) $\forall t \ge 0$ there is a Borel map $p_t : M_F(E) \to M_1\big(M_F(E)\big)$ such that if $\nu \in M_1\big(M_F(E)\big)$ and X satisfies $(LMP)_\nu$, then

$$\mathbb{P}(X_t \in A) = \int p_t(X_0, A)d\nu(X_0) \quad \forall A \in \mathcal{M}_F.$$

Then (a) and (b) of Theorem II.5.1 hold.

Remark. This is a version of the well-known result that uniqueness of the one-dimensional distributions for solutions of a martingale problem implies uniqueness in law and the strong Markov property of any solution. The proof is a simple adaptation of Theorem 4.4.2 of Ethier-Kurtz (1986). Note that (H_1) has already been verified because any limit point in Proposition II.4.2 will satisfy $(LMP)_{\delta_{X_0}}$.

Proof. Let Z satisfy $(LMP)_\nu$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ and T be an a.s. finite $(\mathcal{F}_t)$-stopping time. Choose $A \in \mathcal{F}_T$ so that $\mathbb{P}(A) > 0$ and define Q on $(\Omega, \mathcal{F})$ by $Q(B) = \mathbb{P}\big(B \mid A\big)$. Let $W_t = Z_{T+t}$ and $\mathcal{G}_t = \mathcal{F}_{T+t}$, $t \geq 0$. If $\nu_0 = Q(W_0 \in \cdot)$, we claim W satisfies $(LMP)_{\nu_0}$ on $(\Omega, \mathcal{F}, \mathcal{G}_t, Q)$. Define

$$S_k = \inf\Big\{t : \int_T^{t+T} Z_r(1)dr > k\Big\} \wedge k.$$

One easily checks that $S_k + T$ is an $(\mathcal{F}_t)$-stopping time and S_k is a $(\mathcal{G}_t)$-stopping time. Clearly $S_k \uparrow \infty$ $\mathbb{P}$-a.s., and hence Q-a.s. Let M^Z be the martingale measure associated with Z, and for $\phi \in \mathcal{D}(A)$ let

$$M_t^W(\phi) = W_t(\phi) - W_0(\phi) - \int_0^t W_s\left(A^g\phi\right) ds = M_{T+t}^Z(\phi) - M_T^Z(\phi).$$

Fix $0 \leq s < t$ and $D \in \mathcal{G}_s = \mathcal{F}_{s+T}$. The definition of S_k ensures that

$$N_t = \int_0^t \int 1\big(T \leq s \leq S_k + T\big)\phi(x)M^Z(ds, dx)$$

is an L^2 bounded martingale ($\langle N \rangle_t$ is bounded), and therefore,

$$\begin{aligned} Q\left(\left(M_{t\wedge S_k}^W(\phi) - M_{s\wedge S_k}^W(\phi)\right) 1_D\right) &= \mathbb{P}\left(\left(N_{T+t} - N_{T+s}\right) 1_{D\cap A}\right)/\mathbb{P}(A) \\ &= 0 \end{aligned}$$

by optional sampling, because $D \cap A \in \mathcal{F}_{s+T}$. This proves $M_t^W(\phi)$ is a $(\mathcal{G}_t)$-local martingale under Q and a similar argument shows the same is true of

$$M_t^W(\phi)^2 - \int_0^t W_r(\gamma\phi^2)dr.$$

This shows that W satisfies $(LMP)_{\nu_0}$ on $(\Omega, \mathcal{F}, \mathcal{G}_t, Q)$. (H_2) implies that for $t \geq 0$ and $C \in \mathcal{M}_F$,

$$Q(W_t \in C) = \int p_t(\mu, C)\nu_0(d\mu),$$

that is

$$\mathbb{P}\left(Z_{T+t} \in C \mid A\right) = \mathbb{E}\big(p_t(Z_T, C) \mid A\big),$$

and so

$$(II.5.1) \qquad \mathbb{P}\left(Z_{T+t} \in C \mid \mathcal{F}_T\right) = p_t(Z_T, C) \quad \mathbb{P}\text{-a.s.}$$

Therefore $\{Z_t\}$ is $(\mathcal{F}_t)$-strong Markov with Borel transition kernel p_t and initial law ν, and hence the uniqueness in II.5.1 (a) is proved.

(H_1) and the above allow us to use the above Markov kernel to define the law, $\mathbb{P}_{X_0}$ (on Ω_X) of any solution to $(LMP)_{\delta_{X_0}}$. (II.5.1) implies $X_0 \to \mathbb{P}_{X_0}(A)$ is Borel for finite-dimensional A and hence for all A in $\mathcal{F}_X$. It also implies

$$\mathbb{P}\left(Z_{T+\cdot} \in A \mid \mathcal{F}_T\right)(\omega) = \mathbb{P}_{Z_T(\omega)}(A)$$

first for A finite-dimensional and hence for all A in $\mathcal{F}_X$.

Now consider the "canonical solution" to $(LMP)_\nu$, $X_t(\omega) = \omega_t$, on Ω_X under $\mathbb{P}_\nu(\cdot) = \int \mathbb{P}_{X_0}(\cdot) d\nu(X_0)$. It is easy to check that X solves $(LMP)_\nu$ under $\mathbb{P}_\nu$ for any $\nu \in M_1\big(M_F(E)\big)$. (Note that if $S_k = \inf\{t : \int_0^t X_s(1)ds \geq k\} \wedge k$ then $M_{t\wedge S_k}(\phi)$ is a square integrable martingale under $\mathbb{P}_{X_0}$ and $\mathbb{P}_{X_0}\big(M_{t\wedge S_k}(\phi)^2\big) \leq \|\gamma\phi^2\|_\infty k$ for each $X_0 \in M_F(E)$ and so the same is true under $\mathbb{P}_\nu$.) This proves the existence part of Theorem II.5.1(a). By the above $\big(\Omega_X, \mathcal{F}_X, \mathcal{F}_t^X, X, \mathbb{P}_{X_0}\big)$ is Borel strong Markov, and the proof is complete. ∎

To verify (H_2) we first extend $(LMP)_\nu$ to time dependent functions. Recall that X satisfies $(LMP)_\nu$.

Definition. Let $T > 0$. A function $\phi : [0,T] \times E \to \mathbb{R}$ is in $\mathcal{D}(\vec{A})_T$ iff

(1) For any x in E, $t \to \phi(t,x)$ is absolutely continuous and there is a jointly Borel measurable version of its Radon-Nikodym derivative $\dot{\phi}(t,x) = \frac{\partial\phi}{\partial t}(t,x)$ which is bounded on $[0,T] \times E$ and continuous in x for each $t \in [0,T]$.

(2) For any $t \in [0,T]$, $\phi(t,\cdot) \in \mathcal{D}(A)$ and $A\phi_t$ is bounded on $[0,T] \times E$.

Proposition II.5.7. If $\phi \in D(\vec{A})_T$, then

$$X_t(\phi_t) = X_0(\phi_0) + M_t(\phi) + \int_0^t X_s(\dot{\phi}_s + A^g\phi_s)ds \quad \forall t \in [0,T] \quad \text{a.s.}$$

Proof. Set $t_i^n = i2^{-n}$ and define $\phi^n(t,x) = 2^n \int_{t_{i-1}^n}^{t_i^n} \phi(t,x)dt$ if $t_{i-1}^n \leq t < t_i^n$, $i \geq 1$. Clearly $\phi^n \xrightarrow{bp} \phi$. It is easy to see $\phi^n(t,\cdot) \in \mathcal{D}(A)$ and

$$(II.5.2) \qquad A\phi_t^n(x) = \int_{t_{i-1}^n}^{t_i^n} A\phi_u(x)du\, 2^n \quad \text{if} \quad t_{i-1}^n \leq t < t_i^n.$$

By the (local) martingale problem we have

(II.5.3)

$$X_t\big(\phi_t^n\big) = X_{t_{i-1}^n}(\phi_{t_{i-1}^n}^n) + \int_{t_{i-1}^n}^t X_s\left(A^g\phi_{t_{i-1}^n}^n\right) ds + \int_{t_{i-1}^n}^t \int \phi^n(s,x)dM(s,x), \;\; t \in [t_{i-1}^n, t_i^n).$$

By the continuity of X we have for $i \geq 2$

$$X_{t_{i-1}^n}\left(\phi_{t_{i-1}^n}^n\right) = X_{t_{i-1}^n-}\left(\phi_{t_{i-1}^n-}^n\right) + X_{t_{i-1}^n}\left(\phi_{t_{i-1}^n}^n - \phi_{t_{i-2}^n}^n\right),$$

and so for $i \geq 2$ and $t \in [t_{i-1}^n, t_i^n)$,

$$X_t(\phi_t^n) = X_{t_{i-1}^n-}\left(\phi_{t_{i-1}^n-}^n\right) + X_{t_{i-1}^n}\left(\phi_{t_{i-1}^n}^n - \phi_{t_{i-2}^n}^n\right) + \int_{t_{i-1}^n}^{t} X_s\left(A^g\phi_{t_{i-1}^n}^n\right)ds$$
$$+ \int_{t_{i-1}^n}^{t}\int \phi^n(s,x)dM(s,x).$$

If $t \uparrow t_i^n$, we get a telescoping sum which we may sum over $t_i^n \leq t$ $(i \geq 2)$ and add (II.5.3) with $i = 1$ and $t = t_1^n-$, together with the above expression for $X_t(\phi_t^n) - X_{t_i^n-}(\phi_{t_i^n-})$, where $t \in [t_i^n, t_{i+1}^n)$. If $C_t^n = \sum_{i=1}^{\infty} 1(t_i^n \leq t)X_{t_i^n}\left(\phi_{t_i^n}^n - \phi_{t_{i-1}^n}^n\right)$, we get

$$(II.5.4) \qquad X_t(\phi_t^n) = X_0(\phi_0^n) + C_t^n + \int_0^t X_s\left(A\phi_s^n + g\phi_s^n\right)ds + M_t(\phi^n).$$

Note that if $[t] = [2^n t]2^{-n}$, $[t]^+ = \big([2^n t] + 1\big)2^{-n}$, then

$$\begin{aligned}
C_t^n &= \sum_{i=1}^{[2^n t]} \int_{t_{i-1}^n}^{t_i^n} X_{t_i^n}\left(\phi_{s+2^{-n}} - \phi_s\right)2^n ds \\
&= \sum_{i=1}^{[2^n t]} \int_{t_{i-1}^n}^{t_i^n} \int_s^{s+2^{-n}} X_{t_i^n}(\dot\phi_r)dr\, ds 2^n \\
&= \sum_{i=1}^{[2^n t]} \int_{t_{i-1}^n}^{t_i^n} \int_s^{s+2^{-n}} X_{t_i^n}(\dot\phi_r) - X_r(\dot\phi_r)\, dr ds 2^n + \int_0^{2^{-n}} r X_r(\dot\phi_r)\, dr 2^n \\
&\quad + \int_{2^{-n}}^{[t]} X_r(\dot\phi_r)\, dr + \int_{[t]}^{[t]^+} X_r(\dot\phi_r)\big([t]^+ - r\big)\, dr 2^n.
\end{aligned}$$

The sum of the last three terms approach $\int_0^t X_r(\dot\phi_r)\, dr$ for all $t \geq 0$ a.s. If

$$h_n(r) = \sup\{|X_u(\dot\phi_r) - X_r(\dot\phi_r)| : |u - r| \leq 2^{-n}, u \geq 0\},$$

then $h_n \stackrel{bp}{\to} 0$ a.s. by the continuity of $X_u(\dot\phi_r)$ in u and the first term is at most

$$\int_0^{[t]^+} h_n(r)\, dr \to 0 \qquad \forall t \geq 0 \quad \text{a.s.}$$

We have proved that $C_t^n \to \int_0^t X_r(\dot{\phi}_r)\,dr$ for all $t \geq 0$ a.s. By (II.5.2) we also have

$$\int_0^t X_s(A\phi_s^n)\,ds = \sum_{i=1}^{[2^n t]+1} \int_{t_{i-1}^n}^{t_i^n \wedge t} \int_{t_{i-1}^n}^{t_i^n} X_s(A\phi_u)\,du\,ds 2^n,$$

and an argument very similar to the above shows that

$$\lim_{n\to\infty} \int_0^t X_s(A\phi_s^n)\,ds = \int_0^t X_s(A\phi_s)ds \qquad \forall t \geq 0 \quad \text{a.s.}$$

By considering $\langle M(\phi) - M(\phi^n)\rangle_t$ we see that

$$\sup_{t\leq K} |M_t(\phi) - M_t(\phi^n)| \xrightarrow{\mathbb{P}} 0 \quad \text{as} \quad n \to \infty \quad \text{for all} \quad K > 0.$$

The other terms in (II.5.4) converge trivially by Dominated Convergence and so we may let $n \to \infty$ in (II.5.4) to complete the proof. ∎

Not surprisingly the above extension is also valid for the martingale problem for Y.

Notation. If $\phi \in \mathcal{D}(\vec{A})_T$, let $\vec{A}\phi(t,x) = \dot{\phi}(t,x) + A\phi_t(x)$ for $(t,x) \in [0,T] \times E$.

Proposition II.5.8. If $\phi \in \mathcal{D}(\vec{A})_T$, then

$$N_t = \phi(t, Y_t) - \phi(0, Y_0) - \int_0^t \vec{A}\phi(s, Y_s)\,ds \quad t \in [0,T]$$

is a bounded a.s. cadlag $\mathcal{D}_t$-martingale under P^x for all $x \in E$. Its jumps are contained in the jumps of Y a.s.

Proof. The continuity properties of ϕ imply that

$$\lim_{s\to t+} \phi(s, Y_s) = \phi(t, Y_t) \ \text{ for all } t \in [0,T) \ P^x - a.s.$$

and

$$\lim_{s\to t-} \phi(s, Y_s) = \phi(t, Y_{t-}) \ \text{ for all } t \in (0,T] \ P^x - a.s.$$

Therefore N is a.s. cadlag on $[0,T]$ and can only jump at the jump times of Y a.s. The definition of $\mathcal{D}(\vec{A})_T$ implies that ϕ and $\vec{A}\phi$ are bounded on $[0,T] \times E$ and hence N is also uniformly bounded.

Take mean values in Proposition II.5.7 with $g \equiv 0$ and $X_0 = \delta_x$ and use Remark II.5.5 (2) and Exercise II.5.2 (b)(i) below to see that

$$P_t\phi_t(x) = \phi_0(x) + \int_0^t P_s(\vec{A}\phi_s)(x)\,ds \ \text{ for all } (t,x) \in [0,T] \times E.$$

If $u\in[0,T)$ is fixed then $(t,x)\mapsto\phi(u+t,x)$ is in $\mathcal{D}(\vec{A})_{T-u}$ and so the above implies

$$P_t\phi_{u+t}(x)=\phi_u(x)+\int_0^t P_s(\vec{A}\phi_{s+u})\,ds\ \forall(t,x)\in[0,T-u]\times E\quad\forall u\in[0,T).$$

It is now a simple exercise using the Markov property of Y to see that the above implies that N_t is a $\mathcal{D}_t$-martingale under each P^x. ∎

Green Function Representation

Let $P_t^g\phi(x)=E^x\Big(\phi(Y_t)\exp\Big\{\int\limits_0^t g(Y_s)\,ds\Big\}\Big)$.

Exercise II.5.1. (a) Show that $P_t^g:C_b(E)\to C_b(E)$.
Hint: One approach is to use a Taylor series for exp and recall $P_t:C_b(E)\to C_b(E)$.
(b) Show $\phi\in\mathcal{D}(A)\Leftrightarrow(P_t^g\phi-\phi)t^{-1}\overset{bp}{\to}\psi\in C_b(E)$ as $t\downarrow 0$, and in this case $\psi=A^g\phi$.
(c) Show that $P_t^g:\mathcal{D}(A)\to\mathcal{D}(A)$ and $\frac{d}{dt}P_t^g\phi=A^gP_t^g\phi=P_t^gA^g\phi\quad\forall\phi\in\mathcal{D}(A)$.

The next Exercise will be used extensively.

Exercise II.5.2. Assume X solves $(LMP)_\nu$.
(a) Prove that $\forall\phi\in b\mathcal{E}$

$$(GFR)\qquad X_t(\phi)=X_0\left(P_t^g\phi\right)+\int\limits_0^t\int P_{t-s}^g\phi(x)dM(s,x)\qquad\text{a.s.}\quad\forall t\geq 0.$$

Hint: Assume first $\phi\in\mathcal{D}(A)$ and apply Proposition II.5.7 for an appropriate choice of $\phi(s,x)$.
(b) Assume $\nu=\delta_{X_0}$. (i) Show that $\mathbb{P}\big(X_t(\phi)\big)=X_0\big(P_t^g\phi\big)$ $\forall t\geq 0$ and $\phi\in b\mathcal{E}$.
(ii) Show that if $0\leq s\leq t$ and $\phi,\psi\in b\mathcal{E}$,

$$\mathbb{P}\left(X_s(\phi)X_t(\psi)\right)=X_0\left(P_s^g\phi\right)X_0\left(P_t^g\psi\right)+\int\limits_0^s X_0(P_r^g\left(\gamma P_{s-r}^g\phi P_{t-r}^g\psi\right))\,dr.$$

Hint: Recall Remark II.5.5.

Definition. If W is an $M_F(E)$-valued random vector and $\phi\in b\mathcal{E}_+$, $L_W(\phi)=E(e^{-W(\phi)})\equiv E\big(e_\phi(W)\big)$ is the Laplace functional of W, or of $P_W=P(W\in\cdot)$.

Lemma II.5.9. Assume $D_0\subset(b\mathcal{E})_+$ satisfies $\overline{D_0}^{bp}=b\mathcal{E}_+$. Then $L_W(\phi)=L_{W'}(\phi)$ for all $\phi\in D_0$ implies $P_W=P_{W'}$.

Proof. Clearly equality of L_W and $L_{W'}$ on D_0 implies $L_W=L_{W'}$ on $b\mathcal{E}_+$. An elementary argument (see the proof of Proposition 3.4.4 of Ethier-Kurtz (1986)) shows that there is a countable convergence determining set $V\subset C_b(E)_+$ (i.e. $\nu_n\to\nu$ in $M_F(E)\Leftrightarrow\nu_n(\phi)\to\nu(\phi)\quad\forall\phi\in V$). For any $\phi\in V$, $\nu\to\nu(\phi)$ is measurable with respect to $\sigma(e_\phi:\phi\in V)$. This implies that the class of open sets in $M_F(E)$, and hence $\mathcal{M}_F$, is contained in $\sigma(e_\phi:\phi\in V)$. Apply the Monotone Class Lemma II.5.2 with $\mathcal{H}=\{\Phi\in b\mathcal{M}_F:E\big(\Phi(W)\big)=E\big(\Phi(W')\big)\}$ and $\mathcal{H}_0=\{e_\phi:\phi\in C_b(E)_+\}$ to see that $L_W=L_{W'}$ on $C_b(E)_+$ implies $P_W=P_{W'}$. ∎

We will verify (H_2) by giving an explicit formula for L_{X_t}. If X and X' are independent solutions of $(LMP)_{\delta_{X_0}}$ and $(LMP)_{\delta_{X_0'}}$, respectively, then it is easy to check that $X+X'$ satisfies $(LMP)_{\delta_{X_0+X'_0}}$. This "additive property" is also clear for the approximating branching particle systems considered in Proposition II.4.2; the particles do not interact and so superimposing two such systems gives another such system. This leads to $L_{X_t+X'_t}=L_{X_t}\cdot L_{X'_t}$ and suggests the use of Laplace functionals to characterize the law of X_t. It also explains the "multiplicative" part of the terminology in "critical multiplicative branching measure diffusions", the catchy name for Dawson-Watanabe superprocesses prior to 1987.

Let $\psi\in\mathcal{D}(\vec{A})_t$ for a fixed $t>0$ and $f\in C_b(E)_+$. By Proposition II.5.7 and Itô's Lemma for $u\le t$,

$$\begin{aligned}\exp&\Big\{-X_u(\psi_u)-\int_0^u X_s(f)\,ds\Big\}\\ &=\exp\big(-X_0(\psi_0)\big)-\int_0^u\int\exp\Big\{-X_s(\psi_s)-\int_0^s X_r(f)\,dr\Big\}\psi(s,x)dM(s,x)\\ &\quad+\int_0^u\exp\Big\{-X_s(\psi_s)-\int_0^s X_r(f)\,dr\Big\}\Big[-X_s(\dot\psi_s+A^g\psi_s+f-\gamma\psi_s^2/2)\Big]\,ds.\end{aligned}$$

Let N_u denote the stochastic integral on the righthand side. Let $\phi\in\mathcal{D}(A)_+$. Now choose a non-negative ψ so that the drift term vanishes, and $\psi_t=\phi$, i.e.,

$$(II.5.5)\qquad \dot\psi_s+A^g\psi_s+f-\gamma\psi_s^2/2=0,\quad 0\le s\le t\quad \psi_t=\phi.$$

The previous equation then becomes

$$(II.5.6)\qquad \exp\Big\{-X_u(\phi)-\int_0^u X_s(f)\,ds\Big\}=\exp\big(-X_0(\psi_0)\big)+N_u\quad u\le t.$$

This shows that the local martingale N is bounded and therefore is a martingale satisfying $E(N_t)=0$. Take expectations in (II.5.6) with $u=t$ to see that

$$(II.5.7)\qquad E\Big(\exp\Big\{-X_t(\phi)-\int_0^t X_s(f)\,ds\Big\}\Big)=\int e^{-X_0(\psi_0)}d\nu(X_0).$$

If $a(x,\lambda)=g(x)\lambda-\gamma(x)\lambda^2/2$ and $V_s\equiv V_s^f\phi\equiv\psi_{t-s}$, then $\psi\in\mathcal{D}(\vec{A})_t$ iff $V\in\mathcal{D}(\vec{A})_t$, and (II.5.5) and (II.5.7) (for all $t\ge0$) become, respectively:

$$(SE)_{\phi,f}\qquad \frac{\partial V_s}{\partial s}=AV_s+a(\cdot,V_s)+f,\quad s\ge0,\quad V_0=\phi,$$

and

$$(LE)\qquad E\Big(\exp\Big\{-X_t(\phi)-\int_0^t X_s(f)\,ds\Big\}\Big)=\int e^{-X_0(V_t^f\phi)}d\nu(X_0)\quad\forall t\ge0.$$

These arguments trivially localize to $t\in[0,T]$ and hence we have proved:

Proposition II.5.10. Let $\phi \in \mathcal{D}(A)_+$ and $f \in C_b(E)_+$. If $V \in \mathcal{D}(\vec{A})_T$ for some $T > 0$ and is a non-negative solution V to $(SE)_{\phi,f}$ on $[0,T] \times E$, then (LE) holds for all $0 \le t \le T$ for any X satisfying $(LMP)_\nu$.

The next result will extend this to a larger class of ϕ and f and solutions to the following mild form of (SE):

$$(ME)_{\phi,f} \qquad V_t = P_t\phi + \int_0^t P_{t-s}\big(f + a(\cdot, V_s)\big)\, ds.$$

It is easy to check, e.g. by using Proposition II.5.8 to write $V_{t-s}(Y_s)$ as the sum of a martingale and process of bounded variation, that any solution to (SE) with $V|_{[0,t]\times E} \in D(\vec{A})_t$ for all $t \ge 0$ satisfies (ME). Conversely, a solution to (ME) will satisfy

$$(II.5.8) \qquad \frac{V_{t+h} - V_t}{h} = \frac{(P_h - I)}{h} V_t + \frac{1}{h}\int_0^h P_r\big(a(\cdot, V_{t+h-r}) + f\big)\, dr,$$

and hence should satisfy (SE) provided that these limits exist.

Theorem II.5.11. Let $\phi, f \in b\mathcal{E}_+$.
- (a) There is a unique jointly Borel measurable solution $V_t^f\phi(x)$ of $(ME)_{\phi,f}$ such that $V^f\phi$ is bounded on $[0,T] \times E$ for all $T > 0$. Moreover $V^f\phi \ge 0$.
- (b) If, in addition, $\phi \in \mathcal{D}(A)_+$ and f is continuous, then $V^f\phi\ |_{[0,T]\times E} \in \mathcal{D}(\vec{A})_T$ $\forall T > 0$, $\dot{V}_t^f\phi(x)$ and $AV_t^f(x)$ are continuous in t (as well as x), and $V^f\phi$ satisfies $(SE)_{\phi,f}$.
- (c) If X is any solution of $(LMP)_\nu$ then (LE) holds.

In view of the locally Lipschitz nature of $a(x,\lambda)$ in λ, (a) and (b) of the above result are to be expected and will follow from a standard fixed point argument, although some care is needed as a does not have linear growth in λ. The regularity of the solutions in (b) will be the delicate part of the argument. Note that if the spatial motion is Brownian, the argument here may be simplified considerably because of the regularity of the Brownian transition density.

(c) follows immediately for $\phi \in \mathcal{D}(A)_+$ and $f \in C_b(E)_+$ by Proposition II.5.10. It is then not hard to derive (LE) for all $\phi, f \in b\mathcal{E}_+$ by taking bounded pointwise limits.

We defer the details of the proof until the next section and now complete the **Proof of Theorem II.5.1.** We first verify (H_2) of Theorem II.5.6. If X satisfies $(LMP)_{\delta_{X_0}}$ for $X_0 \in M_F(E)$, then (from Theorem II.5.11) for each $\phi \in \mathcal{D}(A)_+$

$$(II.5.9) \qquad \mathbb{P}\left(e^{-X_t(\phi)}\right) = e^{-X_0(V_t^0\phi)}.$$

This uniquely determines the law, $p_t(X_0, \cdot)$, of X_t by Lemma II.5.9. Let $X_0^n \to X_0$ in $M_F(E)$, then for any $\phi \in C_b(E)$, the mean measure calculation in Exercise II.5.2(b) shows that

$$\int \mu(\phi) p_t(X_0^n, d\mu) = X_0^n(P_t^g\phi) \to X_0(P_t^g\phi) = \int \mu(\phi) p_t(X_0, d\mu).$$

This weak convergence shows that if $\varepsilon > 0$ there is a compact subset of E, K, such that $\sup_n \int \mu(K^c) p_t(X_0^n, d\mu) < \varepsilon$. This shows that $\{p_t(X_0^n, \cdot) : n \in \mathbb{N}\}$ is tight on $M_F(E)$. For example, one can apply Theorem II.4.1 to the set of constant $M_F(E)$-valued processes. (II.5.9) shows that for $\phi \in \mathcal{D}(A)_+$

$$\int e_\phi(\mu) p_t(X_0^n, d\mu) = e^{-X_0^n(V_t^0\phi)} \to e^{-X_0(V_t^0\phi)} = \int e_\phi(\mu) p_t(X_0, d\mu),$$

and hence that $p_t(X_0, \cdot)$ is the only possible weak limit point. We have proved that $X_0 \to p_t(X_0, \cdot)$ is continuous, and in particular is Borel.

If X satisfies $(LMP)_\nu$, then Theorem II.5.11 (c) shows that

$$\mathbb{P}\big(e_\phi(X_t)\big) = \int e^{-X_0(V_t^0\phi)} d\nu(X_0) = \int \int e_\phi(\mu) p_t(X_0, d\mu) d\nu(X_0) \quad \forall \phi \in b\mathcal{E}_+,$$

and so (H_2) follows by Lemma II.5.9. This allows us to apply Theorem II.5.6 and infer (a) and (b). The above continuity of $p_t(X_0, \cdot)$ in X_0 implies (d). Finally the uniqueness in (a) shows that all the weak limit points in Proposition II.4.2 coincide with $\mathbb{P}_{X_0}$ and so the convergence in (c) follows. ∎

The following Feynman-Kac formula shows solutions to $(ME)_{\phi,f}$ for non-negative ϕ and f are necessarily non-negative and will be useful in Chapter III.

Proposition II.5.12. Suppose ϕ, $f \in b\mathcal{E}$ and $V : [0, T] \to \mathbb{R}$ is a bounded Borel function satisfying $(ME)_{\phi,f}$ for $t \le T$. For $u \le t \le T$ define

$$C_u = C_u^{(t)} = \int_0^u g(Y_r) - \frac{\gamma(Y_r)}{2} V(t - r, Y(r))\, dr.$$

Then for all $(t, x) \in [0, T] \times E$,

$$V(t, x) = E^x(\phi(Y_t) e^{C_t}) + \int_0^t E^x(f(Y_r) e^{C_r})\, dr.$$

Proof. Let $0 \le s \le t$. Use $(ME)_{\phi,f}$ with $t - s$ in place of t and apply P_s to both sides to derive

$$\begin{aligned} P_s V_{t-s} &= P_t\phi + \int_0^{t-s} P_{t-r}(f + gV_r - \gamma V_r^2/2)\, dr \\ &= V_t - \int_0^s P_r(f + gV_{t-r} - \gamma V_{t-r}^2/2) dr. \end{aligned}$$

The Markov Property now shows that

$$N_s = V_{t-s}(Y_s) - V_t(Y_0) + \int_0^s f(Y_r) + g(Y_r)V_{t-r}(Y_r) - \gamma(Y_r)V_{t-r}(Y_r)^2/2\, dr, \quad s \le t$$

is a bounded $\mathcal{D}_s$-martingale under P^x for all x. Itô's Lemma then implies

$$e^{C_s} V_{t-s}(Y_s) = V_t(x) - \int_0^s e^{C_r} f(Y_r)\, dr + \int_0^s e^{C_r} dN_r.$$

The stochastic integral is a mean zero L^2-martingale under each P^x and so we may set $s = t$ and take means to complete the proof. ∎

Extinction Probabilities

The Laplace functional equation (LE) is a powerful tool for the analysis of X. As a warm-up we use it to calculate extinction probabilities for X. Assume $g(\cdot) \equiv g \in \mathbb{R}$, $\gamma(\cdot) \equiv \gamma > 0$, and $X_0 \in M_F(E)$ is deterministic. Then, setting $\phi \equiv 1$ in $(LMP)_{\delta_{X_0}}$, we see that $X(1)$ satisfies the martingale problem characterizing the solution of

$$(II.5.10) \qquad X_t(1) = X_0(1) + \int_0^t \sqrt{\gamma X_s(1)} dB_s + \int_0^t g X_s(1)\, ds,$$

where B is a linear Brownian motion. An immediate consequence of Theorem II.5.1(c) is $X^n_\cdot(1) \stackrel{w}{\Rightarrow} X_\cdot(1)$ in $D(\mathbb{R})$, which when $g \equiv 0$ reduces to Feller's Theorem II.1.2.

Assume first $g = 0$ and for $\lambda > 0$ let $V_t = V_t^\lambda$ solve

$$\frac{\partial V_t}{\partial t} = AV_t - \frac{\gamma V_t^2}{2}, \quad V_0 \equiv \lambda.$$

Clearly the solution is independent of x and so one easily gets

$$V_t^\lambda = 2\lambda(2 + \lambda t \gamma)^{-1}.$$

(LE) implies that the Laplace functional of the total mass of the DW-superprocess is

$$(II.5.11) \qquad \mathbb{P}_{X_0}\left(e^{-\lambda X_t(1)}\right) = \exp\left\{-\frac{X_0(1) 2\lambda}{2 + \lambda t \gamma}\right\}.$$

Let $\lambda \to \infty$ to see

$$(II.5.12) \qquad \mathbb{P}_{X_0}(X_t = 0) = \mathbb{P}_{X_0}(X_s = 0\, \forall s \geq t) = \exp\left\{\frac{-2X_0(1)}{t\gamma}\right\}.$$

In particular, by letting $t \to \infty$ we see that X becomes extinct in finite time $\mathbb{P}_{X_0}$-a.s. See Knight (1981, p. 100) for the transition density of this total mass process.

Exercise. II.5.3. Assume $\gamma(\cdot) \equiv \gamma > 0, g(\cdot) \equiv g$ are constants.

(a) Find $\mathbb{P}_{X_0}(X_s \equiv 0 \quad \forall s \geq t)$. $\left(\text{Answer} : \exp\left\{\frac{-2X_0(1)g}{\gamma(1-e^{-gt})}\right\} \text{ if } g \neq 0.\right)$

(b) Show that $\mathbb{P}_{X_0}$ (X becomes extinct in finite time) $= \begin{cases} 1 & \text{if } g \leq 0 \\ \exp\left\{\frac{-2X_0(1)g}{\gamma}\right\} & \text{if } g > 0 \end{cases}$.

(c) If $g > 0$ prove that $\mathbb{P}_{X_0}$-a.s.

$$X \text{ becomes extinct in finite time or } \lim_{t\to\infty} X_t(1) = \infty.$$

Hint. Show that $e^{-\lambda X_t(1)}$ is a supermartingale for sufficiently small $\lambda > 0$ and $\lim_{t\to\infty} \mathbb{E}_{X_0}\left(e^{-\lambda X_t(1)}\right) = \exp\left\{\frac{-2X_0(1)g}{\gamma}\right\}$.

Exercise II.5.4, Assume $X_0 \in M_F(E) - \{0\}$, $\nu_N(x, dk) \equiv \nu_N(dk)$ is independent of x and $g_N \equiv 0$. Prove that for ϕ bounded and measurable,

$$E(X_t^N(\phi)|X_s^N(1), s \geq 0) = \frac{X_0(P_t\phi)}{X_0(1)} X_t^N(1).$$

Conclude that if X is the $(Y, \gamma, 0)$-DW-superprocess ($\gamma \geq 0$ is constant), then

$$E(X_t(\phi)|X_s(1), s \geq 0) = \frac{X_0(P_t\phi)}{X_0(1)} X_t(1).$$

Hint. Condition first on the larger σ-field $\sigma(N^\alpha, \alpha \in I) \vee \sigma(M_N)$ (recall M_N is the Poisson number of initial points).

Remark II.5.13. Assume that X satisfies $(LMP)_\nu$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where A is the generator of the d-dimensional symmetric stable process of index $\alpha \in (0, 2]$ and the smaller class $C_b^\infty(\mathbb{R}^d)$ is used in place of $\mathcal{D}(A)$. Recalling that $C_b^\infty(\mathbb{R}^d)$ is a core for $\mathcal{D}(A)$ (see Example II.2.4), we may pass to the bounded pointwise closure of $\{(\phi, A\phi) : \phi \in C_b^\infty(\mathbb{R}^d)\}$ in $(LMP)_\nu$ by Dominated Convergence. Here note that if $T_k = \inf\{t : X_t(1) > k\} \wedge k$ and $\phi_n \xrightarrow{bp} \phi$, then $M_{t \wedge T_k}(\phi_n)$ is a bounded martingale for all n. Therefore X satisfies $(LMP)_\nu$ and so is an $(\mathcal{F})_t$-(A, γ, g)-DW superprocess with initial law ν.

Exercise II.5.5. Let X be the $(Y, \gamma, 0)$-DW superprocess starting at $X_0 \in M_F(E)$, where Y is the d-dimensional symmetric stable process of index $\alpha \in (0, 2]$ and $\gamma > 0$ is constant. For $\gamma_0, \lambda > 0$, let $\phi_\lambda(x) = \phi(x\lambda^{-1/\alpha})$ and

$$X_t^{(\lambda)}(\phi) = \frac{\gamma_0}{\lambda} X_{\lambda t}(\phi_\lambda), \quad t \geq 0.$$

Prove that $X^{(\lambda)}$ is a $(Y, \gamma\gamma_0, 0)$-DW superprocess (starting of course at $X_0^{(\lambda)}$).

6. Proof of Theorem II.5.11.

Step 1. If ϕ, $f \in b\mathcal{E}$, there is a $t_{\max} \in (0, \infty]$ and a unique solution V to $(ME)_{\phi,f}$ on $[0, t_{\max})$ which is bounded on $[0, T] \times E$ for all $T < t_{\max}$ and satisfies $\overline{\lim\limits_{t \uparrow t_{\max}}} \|V_t\| = \infty$ if $t_{\max} < \infty$. If in addition ϕ, $f \in C_b(E)$ and $\|P_t\phi - \phi\| \to 0$ as $t \downarrow 0$ (as is the case for $\phi \in \mathcal{D}(A)$), then $V : [0, t_{\max}) \to C_b(E)$ is continuous in norm.

This is a standard fixed point argument which only requires $a(\cdot, 0) = 0$ and

$$\forall K > 0, a\,|_{E \times [-K,K]} \in C_b\big(E \times [-K, K]\big) \text{ and } |a(x, \lambda) - a(x, \lambda')| \leq C_K |\lambda - \lambda'| \tag{II.6.1}$$
$$\text{for all } x \text{ in } E, \lambda, \lambda' \in [-K, K] \text{ for some increasing } \{C_K\} \text{ and } C_0 \geq 1.$$

We start with ϕ, f as in the second part of the above assertion. We will view f as fixed but will let ϕ vary and will choose $\delta = \delta(\|\phi\|) > 0$ below. Define $\psi : C\big([0, \delta], C_b(E)\big) \to C\big([0, \delta], C_b(E)\big)$ by

$$\psi(U)(t) = P_t\phi + \int_0^t P_{t-s}\big(f + a(U_s)\big) ds.$$

Note that $a(U_s) \equiv a(\cdot, U_s(\cdot)) \in C_b(E)$. To see that ψ does map into the above space, note first that $t \to P_t\phi$ is in $C\big(\mathbb{R}_+, C_b(E)\big)$ by our choice of ϕ and the semigroup property. If $0 \le t < t+h \le \delta$, then

$$\begin{aligned}
&\left|\int_0^{t+h} P_{t+h-s}\big(f + a(U_s)\big)ds - \int_0^t P_{t-s}\big(f + a(U_s)\big)ds\right| \\
&\quad \le \left|\int_0^h P_{t+h-s}\big(f + a(U_s)\big)ds\right| + \int_0^t \|P_{t-s}\big(a(U_{s+h})\big) - P_{t-s}\big(a(U_s)\big)\|ds \\
&\quad \le \left[\|f\| + \sup_{s\le\delta}\|a(U_s)\|\right]h + \int_0^t \|a(U_{s+h}) - a(U_s)\|ds \\
&\quad \to 0 \quad \text{as} \quad h \downarrow 0
\end{aligned}$$

by (II.6.1) and $\sup\limits_{s\le\delta}\|U_s\| < \infty$. Hence ψ is as claimed above.

Take $K = 2\|\phi\| + 1$ and for $U \in C\big([0,\delta], C_b(E)\big)$ define $\|U\| = \sup_{t\le\delta}\|U_t\|$ and let $\overline{B}(0,K)$ be the set of such U with $\|U\| \le K$. If $U \in \overline{B}(0,K)$ and

$$0 < \delta \le \varepsilon_1(\|\phi\|) = \big(K - \|\phi\|\big)/\big(\|f\| + KC_K\big),$$

then

$$\|\psi(U)\| \le \|\phi\| + \delta\|f\| + \delta \sup_{|\lambda|\le K}\|a(\cdot,\lambda)\| \le \|\phi\| + \delta[\|f\| + KC_K] \le K$$

and therefore $\psi : \overline{B}(0,K) \to \overline{B}(0,K)$. If, in addition, $0 < \delta \le \varepsilon_2(\|\phi\|) = 1/2C_K$, then an application of (6.1) shows that for $U, V \in \overline{B}(0,K)$

$$\|\psi(U) - \psi(V)\| \le \int_0^\delta \|a(U_s) - a(V_s)\|ds \le C_K\delta\|U - V\| \le \frac{1}{2}\|U - V\|.$$

Now let $\delta = \delta(\|\phi\|) = \min\big(\varepsilon_1(\|\phi\|), \varepsilon_2(\|\phi\|)\big)$ and note that

$$(II.6.2) \qquad \inf_{0\le r\le M}\delta(r) > 0 \quad \text{for any} \quad M > 0.$$

Then ψ is a contraction on the complete metric space $\overline{B}(0,K)$ and so has a unique fixed point V_t which solves $(ME)_{\phi,f}$ for $t \le \delta$.

To repeat this construction with V_δ in place of ϕ we must check that $\|P_hV_\delta - V_\delta\| \to 0$ as $h \downarrow 0$. Use $(ME)_{\phi,f}$ at $t = \delta$ to see this reduces to

$$\|\int_0^\delta P_{h+\delta-s}a(V_s) - P_{\delta-s}a(V_s)ds\| \to 0 \quad \text{as} \quad h \downarrow 0.$$

The above norm is at most $(0 < h < \delta)$

$$\begin{aligned}&\|\int_0^h P_{h+\delta-s}a(V_s)ds\| + \|\int_{\delta-h}^{\delta} P_{\delta-s}a(V_s)ds\| + \int_0^{\delta-h} \|P_{\delta-s}\big(a(V_{s+h}) - a(V_s)\big)\|ds\\ &\quad\le 2\sup_{s\le\delta}\|a(V_s)\|h + \int_0^{\delta-h}\|a(V_{s+h}) - a(V_s)\|ds\\ &\quad\to 0 \quad \text{as} \quad h\downarrow 0\end{aligned}$$

by the norm-continuity of V and (II.6.1). By repeating the previous argument with V_δ in place of ϕ we can extend $V.$ to a norm-continuous solution to $(ME)_{\phi,f}$ on $[0,\delta_1+\delta_2]$ where $\delta_1 = \delta(\|\phi\|)$ and $\delta_2 = \delta(\|V_{\delta_1}\|)$. Continue inductively to construct a norm-continuous solution to $(ME)_{\phi,f}$ on $[0,t_{\max})$, where $t_{\max} = \sum\limits_{n\ge1}^{\infty}\delta_n$ and $\delta_{n+1} = \delta(\|V_{\delta_1+\cdots+\delta_n}\|)$. If $t_{\max}<\infty$, clearly $\lim\delta_n = 0$ and so (II.6.2) implies $\lim\limits_{n\to\infty}\|V_{\delta_1+\cdots+\delta_{n-1}}\| = \infty$ and hence $\overline{\lim\limits_{t\uparrow t_{\max}}}\|V_t\| = \infty$.

For $\phi,f\in b\mathcal{E}$ one can use the same existence proof with $L^\infty\big([0,\delta]\times E\big)$, the space of bounded Borel functions with the supremum norm, in place of $C_b\big([0,\delta],C_b(E)\big)$. We need the fact $P_t\phi(x)$ is jointly Borel which is clear for $\phi\in C_b(E)$ (because (II.2.2) and (QLC) imply continuity in each variable separately) and hence for all $\phi\in b\mathcal{E}$ by a monotone class argument. It follows easily that $\psi(U)(t,x)$ is Borel and the argument proceeds as above to give a Borel solution of $(ME)_{\phi,f}$ on $[0,t_{\max})\times E$.

Turning to uniqueness, assume V and $\tilde V$ are solutions to $(ME)_{\phi,f}$ on $[0,t_{\max})$ and $[0,\tilde t_{\max})$, respectively so that V and $\tilde V$ are locally (in t) bounded and, in particular, $K = \sup\limits_{s\le t}\|V_s\|\vee\|\tilde V_s\| < \infty$ for $t < t_{\max}\wedge\tilde t_{\max}$. Then for such a t and K,

$$\|V_t - \tilde V_t\| \le C_K\int_0^t \|V_s-\tilde V_s\|ds$$

which implies $V=\tilde V$ on $[0,t_{\max}\wedge\tilde t_{\max})$ by Gronwall's Lemma ($s\to\|V_s-\overline{V}_s\|$ is universally measurable). Clearly $t_{\max}<\tilde t_{\max}$ is impossible because then $\overline{\lim\limits_{t\uparrow t_{\max}}}\|V_t\| = \infty$ would imply $\lim\limits_{t\uparrow t_{\max}}\|\tilde V_t\| = \infty$ which is impossible for $t_{\max}<\tilde t_{\max}$ by our local boundedness assumption on the solution $\tilde V$. Therefore $t_{\max} = \tilde t_{\max}$ and so $V=\tilde V$. This completes Step 1.

Step 2. If $\phi\in\mathcal{D}(A)$ and $f\in C_b(E)$, then the above solution satisfies the conclusions of (b) for $T,t<t_{\max}$.

The key step will be the existence of $\frac{\partial V}{\partial t}$. In addition to (II.6.1), the only property of a we will use is

$$(II.6.3)\qquad \begin{aligned}&a'(x,\lambda)\equiv\frac{\partial}{\partial\lambda}a(x,\lambda)\in C(E\times\mathbb{R}_+) \text{ and satisfies}\\ &\lim_{\delta\downarrow0}\sup_{|\lambda|\le K}\|a'(\lambda)-a'(\lambda+\delta)\| = 0 \text{ and } \sup_{|\lambda|\le K}\|a'(\cdot,\lambda)\|<\infty\quad\forall K>0.\end{aligned}$$

Fix $0 < T < t_{\max}$. Recall from (II.5.8) that for $h > 0$, if

$$R_t^h = h^{-1} \int_0^h P_r\big(f + a(V_{t+h-r})\big) dr,$$

then $V_t^h = (V_{t+h} - V_t)h^{-1}$ satisfies

$$(II.6.4) \qquad V_t^h = \frac{(P_h - I)}{h} V_t + R_t^h.$$

The norm continuity of V_t (from Step 1) and (II.6.1) show that as $h, r \downarrow 0$ $(r < h)$, $\|a(V_{t+h-r}) - a(V_t)\| \to 0$ and so $P_r\big(f + a(V_{t+h-r})\big) \xrightarrow{bp} f + a(V_t)$ on $[0,T] \times E$. It follows that $R_t^h \xrightarrow{bp} f + a(V_t) \in C_b([0,T] \times E)$ as $h \to 0+$. Therefore it is clear from (II.6.4) that if

$$(II.6.5) \qquad V_t^h \xrightarrow{bp} \dot{V}_t \quad \text{on} \quad [0,T] \times E \quad \text{and the limit is continuous in each variable separately,}$$

then the conclusions of (b) hold on $[0,T] \times E$ and Step 2 is complete.

To prove (II.6.5), write

$$\begin{aligned} V_t^h = {} & \frac{P_{t+h}\phi - P_t\phi}{h} + h^{-1} \int_0^h P_{t+h-s}\big(f + a(V_s)\big) ds \\ & + \int_0^t P_{t-s}\left(\big(a(V_{s+h}) - a(V_s)\big)h^{-1} - a'(V_s)V_s^h\right) ds + \int_0^t P_{t-s}\big(a'(V_s)V_s^h\big) ds \end{aligned}$$

$$(II.6.6) \equiv 1_h + 2_h + 3_h + 4_h.$$

The norm continuity of V_s, and hence of $a(V_s)$, together with $\phi \in \mathcal{D}(A)$ imply

$$(II.6.7) \qquad 1_h + 2_h \xrightarrow{bp} P_t(A\phi) + P_t\big(f + a(\phi)\big) \quad \text{on} \quad [0,T] \times E \quad \text{as} \quad h \downarrow 0.$$

Note also that the limit is continuous in each variable if the other is fixed. The mean value theorem shows there is a $\zeta_s^h(x)$ between $V_s(x)$ and $V_{s+h}(x)$ such that

$$\left[\big(a(V_{s+h}) - a(V_s)\big)h^{-1} - a'(V_s)V_s^h\right](x) = \left(a'\left(x, \zeta_s^h(x)\right) - a'\left(x, V_s(x)\right)\right) V_s^h(x).$$

This together with (II.6.3) and the norm continuity of V_s imply

$$(II.6.8) \qquad \sup_x |3_h| \le \eta_h \int_0^t \|V_s^h\| ds \quad \text{for some} \quad \eta_h \to 0 \quad \text{as} \quad h \to 0+.$$

Our local boundedness condition on a' (see (II.6.3)) and norm continuity of V imply

$$\sup_x |4_h| \le C \int_0^t \|V_s^h\| ds.$$

Use the above bounds in (II.6.6) to get

$$\|V_t^h\| \le C + C\int_0^t \|V_s^h\| ds, \quad t \le T$$

and hence

$$\sup_{t\le T} \|V_t^h\| \le Ce^{CT}. \tag{II.6.9}$$

We now may conclude from (II.6.8) that

$$\sup_{t\le T,x} |3_h| \to 0 \quad \text{as} \quad h\downarrow 0. \tag{II.6.10}$$

The above results and (II.6.6) suggest that $\dot V_t$ (if it exists) should solve

$$W_t = P_t\big(A\phi + f + a(\phi)\big) + \int_0^t P_{t-s}\big(a'(V_s)W_s\big)ds. \tag{II.6.11}$$

A slight modification of Step 1 shows there is a unique solution of (II.6.11) in $L^\infty\big([0,T], C_b(E)\big)$. To see this, set $\theta = A\phi + f + a(\phi) \in C_b(E)$ and define

$$\psi : L^\infty([0,T], C_b(E)) \to L^\infty([0,T], C_b(E))$$

by

$$\psi(W)(t) = P_t\theta + \int_0^b P_{t-s}\big(a'(V_s)W_s\big)ds.$$

Clearly $h(s,x,\lambda) = a'\big(x, V_s(1)\big)\lambda$ is Lipschitz in λ uniformly in $(s,x)\in[0,T]\times E$ and so as in Step 1 we get the existence of a unique fixed point W first on $L^\infty\big([0,\delta], C_b(E)\big)$ for appropriate $\delta > 0$ and then on $L^\infty\big([0,T], C_b(E)\big)$ by iteration because the linear growth of h in λ means the solution cannot explode. As $W_t(x)$ is continuous in x for each t and continuous in t for each x (see (II.6.11)), to prove (II.6.5) it suffices to show

$$V_t^h \overset{bp}{\to} W_t \quad \text{on} \quad [0,T]\times E. \tag{II.6.12}$$

In view of (II.6.9) we only need establish pointwise convergence. For this we may fix $h_n \downarrow 0$ and define $r(t,x) = \overline{\lim_{n\to\infty}} |V_t^{h_n}(x) - W_t(x)|$ which is bounded on $[0,T]\times E$ because W is. Apply (II.6.6), (II.6.7), (II.6.10) and (II.6.11) to see that

$$\begin{aligned} r(t,x) &= \overline{\lim_{n\to\infty}} \left| \int_0^t P_{t-s}\left(a'(V_s)(V_s^{h_n} - W_s)\right)(x)ds \right| \\ &\le C\int_0^t P_{t-s}(r_s)(x)ds, \end{aligned}$$

and so

$$\|r_t\| \le C \int_0^t \|r_s\| ds.$$

This implies $r \equiv 0$ and hence (II.6.12). The proof of Step 2 is complete.

Step 3. If $\phi, f \in b\mathcal{E}_+$, then $t_{\max} = \infty$, $V_t = V_t^f \phi \ge 0$ and is bounded on $[0,T] \times E$ $\forall T > 0$ and (LE) holds if X is any solution of $(LMP)_\nu$.

The non-negativity is immediate from Proposition II.5.12. For the other assertions assume first $\phi \in \mathcal{D}(A)_+$, $f \in C_b(E)_+$. Step 2, the non-negativity of $V^f\phi$, and Proposition II.5.10 show that (LE) is valid for $t < t_{\max}$. If $\overline{g} = \sup_x g(x)$, $(ME)_{\phi,f}$ and the non-negativity of $V^f\phi$ imply

$$\|V_t^f \phi\| \le \|\phi\| + t\|f\| + \overline{g} \int_0^t \|V_s^f \phi\| ds, \quad t < t_{\max}$$

and therefore

$$(II.6.14) \qquad \|V_t^f \phi\| \le (\|\phi\| + t\|f\|)\, e^{\overline{g}t}, \quad t < t_{\max}.$$

This means $\|V_t^f \phi\|$ cannot explode at a finite $t_{\max}$ and so $t_{\max} = \infty$.

Turning to more general (ϕ, f), let

$$\mathcal{H} = \left\{ (\phi, f) \in (b\mathcal{E}_+)^2 : t_{\max} = \infty, \quad \text{(LE) holds} \right\}.$$

Assume $(\phi_n, f_n) \stackrel{bp}{\to} (\phi, f)$ and $(\phi_n, f_n) \in \mathcal{H}$. By (II.6.14) we have

$$\sup_{n, t \le T} \|V_t^{f_n} \phi_n\| < \infty \quad \forall T > 0.$$

Apply (LE) with $\nu = \delta_{\delta_x}$ and (ϕ_n, f_n) to see that $V_t^{f_n}\phi_n \stackrel{bp}{\to} V_t^\infty$ on $[0,T] \times E$ $\forall T > 0$ (the boundedness is immediate from the above). Now let $n \to \infty$ in $(ME)_{\phi_n, f_n}$ and use Dominated Convergence to see that $V_t^\infty = V_t^f \phi$ and for (ϕ, f), $t_{\max} = \infty$, and (LE) holds by taking limits in this equation for (ϕ_n, f_n). This shows $\mathcal{H}$ is closed under $\stackrel{bp}{\to}$. As $\mathcal{H} \supset \mathcal{D}(A)_+ \times C_b(E)_+$ (by the previous argument) and $\mathcal{D}(A)_+$ is bp-dense in $b\mathcal{E}_+$ (Corollary II.2.3) we may conclude that $\mathcal{H} = (b\mathcal{E}_+)^2$. This completes Step 3 because the boundedness claim is immediate from $t_{\max} = \infty$ and the local boundedness established in Step 1.

(a) is immediate from Steps 1 and 3. (c) follows from Step 3. (b) is clear from Step 2 and $t_{\max} = \infty$ in Step 3. ∎

7. Canonical Measures

Definition. A random finite measure, X, on E is infinitely divisible iff for any natural number n there are i.i.d. random measures $\{X_i : i \le n\}$ such that X and $X_1 + \ldots + X_n$ have the same law on $M_F(E)$.

Example. Let $(X_t, t \ge 0)$ be a (Y, γ, g)-DW-superprocess starting at $X_0 \in M_F(E)$. If $\{X^i_\cdot : i \le n\}$ are iid copies of the above DW-process but starting at X_0/n, then

$$(II.7.1) \qquad X_\cdot \overset{\mathcal{D}}{\equiv} X^1_\cdot + \ldots + X^n_\cdot \quad \text{as continuous} \quad M_F(E)\text{-valued processes.}$$

This follows from Theorem II.5.1, by noting that $X^1 + \ldots + X^n$ satisfies the martingale problem which characterizes the law of X (or by using the convergence theorem and the corresponding decomposition for the approximating branching particle systems). In particular for each fixed $t \ge 0$, X_t is an infinitely divisible random measure.

For our purposes, Chapter 3 of Dawson (1992) is a good reference for infinitely divisible random measures on a Polish space (see also Kallenberg (1983) for the locally compact case). The following canonical representation is essentially derived in Theorem 3.4.1 of Dawson (1992).

Theorem II.7.1. Let X be an infinitely divisible random measure on E such that $E\big(X(1)\big) < \infty$. There is a unique pair (M, R) such that $M \in M_F(E)$, R is a measure on $M_F(E) - \{0\}$ satisfying $\int \nu(1) R(d\nu) < \infty$, and

$$(II.7.2) \quad E\big(\exp\big(-X(\phi)\big)\big) = \exp\Big\{ -M(\phi) - \int 1 - e^{-\nu(\phi)} R(d\nu)\Big\} \ \forall \phi \in (b\mathcal{E})_+.$$

Conversely if M and R are as above, then the right-hand side of (II.7.2) is the Laplace functional of an infinitely divisible random measure X satisfying $E\big(X(1)\big) < \infty$.

Definition. The measure R in (II.7.2) is the canonical measure associated with X.

We will give an independent construction of the canonical measure associated with X_t, a DW-superprocess evaluated at t, below (see Theorem II.7.2 and Exercise II.7.1). There are some slight differences between the above and Theorem 3.4.1 of Dawson (1992) and so we point out the necessary changes in the

Proof of Theorem II.7.1. A measure, μ, on E is locally finite $\big(\mu \in M_{LF}(E)\big)$ iff it is finite on bounded sets. Suppose X is infinitely divisible and $E\big(X(1)\big) < \infty$. Theorem 3.4.1 of Dawson (1992) shows there is a locally finite measure, M, on E and a measure, R, on $M_{LF}(E) - \{0\}$ such that (II.7.2) holds for all $\phi \in (b\mathcal{E})_+$ with bounded support. Fix such a ϕ. Then for $\lambda > 0$

$$(II.7.3) \qquad u(\lambda\phi) \equiv -\log E\left(e^{-X(\lambda\phi)}\right) = \lambda M(\phi) + \int 1 - e^{-\lambda\nu(\phi)} R(d\nu).$$

Since $\nu(\phi)e^{-\lambda\nu(\phi)} \le C_\lambda(1 - e^{-\lambda\nu(\phi)})$ and $\int 1 - e^{-\lambda\nu(\phi)} R(d\nu) < \infty$ for $\lambda > 0$, it follows that $\int \nu(\phi)e^{-\lambda\nu(\phi)}\ R(d\nu) < \infty$ for $\lambda > 0$. An application of the Mean Value and Dominated Convergence Theorems allows us to differentiate (II.7.3) with respect to $\lambda > 0$ and conclude

$$E\left(X(\phi)e^{-\lambda X(\phi)}\right)\left[E\left(e^{-\lambda X(\phi)}\right)\right]^{-1} = M(\phi) + \int \nu(\phi)e^{-\lambda\nu(\phi)} R(d\nu).$$

Let $\lambda \to 0+$ and use Monotone Convergence to see

$$(II.7.4) \qquad E\big(X(\phi)\big) = M(\phi) + \int \nu(\phi)R(d\nu)$$

first for ϕ as above and then for all non-negative measurable ϕ by Monotone Convergence. Take $\phi = 1$ to see M is finite, $\int \nu(1)R(d\nu) < \infty$ and so R is supported by $M_F(E) - \{0\}$. We can also take monotone limits to see that (II.7.2) holds for all $\phi \in (b\mathcal{E})_+$.

For uniqueness note from (II.7.3) that for any ϕ in $(b\mathcal{E})_+$

$$\lim_{\lambda\to\infty} u(\lambda\phi)\lambda^{-1} = M(\phi) + \lim_{\lambda\to\infty} \int \left(1 - e^{-\lambda\nu(\phi)}\right) \lambda^{-1} R(d\nu) = M(\phi),$$

where in the last line we used $(1-e^{-\lambda\nu(\phi)})\lambda^{-1} \le \nu(\phi)$ and Dominated Convergence. This shows that M and $\int h(\nu)R(d\nu)$ are determined by the law of X for h in

$$\mathcal{C} = \{h(\nu) = \sum_1^K b_i e^{-\nu(\phi_i)} : b_i \in \mathbb{R},\ \phi_i \in (b\mathcal{E})_+,\ h(0) = 0\}.$$

Note that for integration purposes $h(\nu) = -\sum\limits_1^K b_i(1 - e^{-\langle \nu,\phi_i\rangle})$, and $\mathcal{C}$ is a vector space closed under multiplication. As in Lemma II.5.9, the Monotone Class Lemma 5.2 shows these integrals determine R.

Assume conversely that (II.7.2) holds for some M, R as in the Theorem. As in Theorem 3.4.1 of Dawson the right-hand side is the Laplace functional of some random measure which clearly must then be infinitely divisible. One then obtains (II.7.4) as above and this shows $E\big(X(1)\big) < \infty$. ∎

Assume now that X is the (Y,γ,g)-DW superprocess with $\gamma(\cdot) \equiv \gamma > 0$ constant, $g \equiv 0$, and law $\mathbb{P}_{X_0}$ if X starts at $X_0 \in M_F(E)$. Let $x_0 \in E$ and consider the approximating branching particle systems, $X^N_\cdot$, in Theorem II.5.1 starting at δ_{x_0} (under $P^N_{\delta_{x_0}}$) and δ_{x_0}/N (under $P^N_{\delta_{x_0}/N}$), and with $g_N \equiv 0$ and $\nu_N(x,dk) = \nu(dk)$ independent of (x,N). In the former case we start N particles at x_0 (see Remark II.3.2) and in the latter we start a single particle at x_0. Let $\phi \in C_b(E)_+$ and write $V_t\phi$ for $V_t^0\phi$, the unique solution of $(ME)_{\phi,0}$. Lemma II.3.3 and Remark II.3.4 (the arguments go through unchanged for our slightly different initial conditions) show that

$$(II.7.5) \qquad NP^N_{\delta_{x_0}/N}\left(X^N_t(\phi)\right) = P^N_{\delta_{x_0}}\left(X^N_t(\phi)\right) = P_t\phi(x_0).$$

Theorem II.5.1 and (LE) imply that

$$\left[P^N_{\delta_{x_0}/N}\big(\exp\left(-X^N_t(\phi)\right)\big)\right]^N = P^N_{\delta_{x_0}}\left(\exp\left(-X^N_t(\phi)\right)\right) \to \exp\left(-V_t\phi(x_0)\right) \text{ as } N \to \infty.$$

Take logarithms and use $\log z \sim z-1$ as $z \to 1$ (the expression under the Nth power must approach 1) to conclude

$$(II.7.6) \qquad \lim_{N\to\infty} \int \left(1 - e^{-X^N_t(\phi)}\right) N dP^N_{\delta_{x_0}/N} = V_t\phi(x_0).$$

Also note by Kolmogorov's Theorem (II.1.1(a)) that

$$(II.7.7) \qquad \lim_{N\to\infty} NP^N_{\delta_{x_0}/N}\left(X^N_t \neq 0\right) = 2/\gamma t.$$

(II.7.7) and (II.7.5) easily imply tightness of $NP^N_{\delta_{x_0}/N}(X^N_t \in \cdot, X^N_t \neq 0)$ and (II.7.6) shows the limit points coincide. The details are provided below.

Theorem II.7.2. For each $x_0 \in E$ and $t > 0$ there is a finite measure $R_t(x_0, \cdot)$ on $M_F(E) - \{0\}$ such that

(i) $NP^N_{\delta_{x_0}/N}(X^N_t \in \cdot, X^N_t \neq 0) \overset{w}{\Rightarrow} R_t(x_0, \cdot)$ on $M_F(E)$ and $x_0 \mapsto R_t(x_0, \cdot)$ is Borel measurable,

(ii) $\mathbb{P}_{X_0}\left(\exp\left(-X_t(\phi)\right)\right) = \exp\left\{-\int\int 1 - e^{-\nu(\phi)} R_t(x_0, d\nu) dX_0(x_0)\right\} \ \forall \phi \in b\mathcal{E}_+$,

(iii) $R_t\left(x_0, M_F(E) - \{0\}\right) = 2/\gamma t, \quad \int \nu(\phi) R_t(x_0, d\nu) = P_t\phi(x_0) \ \forall \phi \in b\mathcal{E}$,

$$R_t\left(x_0, \{\nu : \nu(1) \in A\}\right) = (2/\gamma t)^2 \int 1_A(x) \exp\{-2x/\gamma t\} dx \ \forall A \in \mathcal{B}\big((0,\infty)\big),$$

(iv) $\int \psi(\nu(1))\nu(\phi) R_t(x_0, d\nu) = \int_0^t \psi(\gamma t z/2) z e^{-z}\, dz\, P_t\phi(x_0) \quad \forall \phi \in b\mathcal{E},\ \psi \in b\mathcal{B}(\mathbb{R}_+)$.

Proof. A sequence $\{\mu_N\}$ of finite, non-zero measures on $M_F(E)$ is tight if $\sup_N \mu_N(1) < \infty$ and for any $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset E$ such that $\mu_N\big(\{\nu : \nu(K^c_\varepsilon) > \varepsilon\}\big)/\mu_N(1) < \varepsilon$. For example, one may apply Theorem II.4.1 to the set of constant paths $\tilde{X}^N \equiv \tilde{X}^N_0$ with law $\mu_N/\mu_N(1)$. (II.7.5) and (II.7.7) easily imply these conditions for $\mu_N(\cdot) = NP^N_{\delta_{x_0}/N}(X^N_t \in \cdot, X^N_t \neq 0)$. Let μ_∞ be any weak limit point in $M_F\big(M_F(E)\big)$. Then (II.7.6) implies

$$(II.7.8) \qquad \int 1 - e^{-\nu(\phi)} d\mu_\infty(\nu) = V_t\phi(x_0) \qquad \forall \phi \in C_b(E)_+.$$

Take $\phi \equiv \lambda > 0$ in the above, recall $V_t\lambda = 2\lambda(2 + \lambda t\gamma)^{-1}$ and let $\lambda \to \infty$ to see that

$$\begin{aligned} \mu_\infty\big(M_F(E) - \{0\}\big) &= 2/\gamma t \\ &= \lim_{N\to\infty} \mu_N(1) \qquad \text{(by (II.7.7))} \\ &= \mu_\infty(1). \end{aligned}$$

This shows $\mu_\infty(\{0\}) = 0$ and, together with (II.7.8), implies

$$(II.7.9) \qquad \int e^{-\nu(\phi)} d\mu_\infty(\nu) = 2/\gamma t - V_t\phi(x_0) \qquad \forall \phi \in C_b(E)_+.$$

As in Lemma II.5.9, this uniquely determines μ_∞ and shows $\mu_N \overset{w}{\Rightarrow} \mu_\infty$. The Borel measurability of $R_t(x_0)$ in x_0 is then clear from the Borel measurability of the approximating measures. The proof of (i), (ii) (by (LE)), and the first assertion in (iii) is complete. The second assertion in (iii) is a special case of (iv), proved below. The final assertion in (iii) is obtained by setting $\phi \equiv \lambda$ in (II.7.9), as was already done in the above.

For (iv) it suffices to consider ψ and ϕ bounded and continuous. For the branching particle system described above, $\mathcal{N} = \sigma(N^\alpha : \alpha \in I)$ is independent of $\sigma(Y^\alpha : \alpha \in I)$ and so

$$\begin{aligned} &NP^N_{\delta_{x_0}/N}(\psi(X^N_t(1))X^N_t(\phi)1(X^N_t(1) \neq 0)) \\ &= NP^N_{\delta_{x_0}/N}\big(\psi(X^N_t(1))1(X^N_t(1) \neq 0)\frac{1}{N}\sum_{\alpha\sim t} P^N_{\delta_{x_0}/N}(\phi(Y^\alpha_t)|\mathcal{N})\big) \\ &= NP^N_{\delta_{x_0}/N}(\psi(X^N_t(1))1(X^N_t(1) \neq 0)X^N_t(1))P_t\phi(x_0). \end{aligned}$$

Now let $N \to \infty$ in the above. Lemma II.4.6 and $NP^N_{\delta_{x_0}/N}\left(X^N_t(1)^2\right) \le P^N_{\delta_{x_0}}\left(X^N_t(1)^2\right)$ give us the necessary uniform integrability to use (i) and conclude that

$$\int \psi(\nu(1))\nu(\phi)R_t(x_0, d\nu) = \int \psi(\nu(1))\nu(1)R_t(x_0, d\nu)P_t\phi(x_0),$$

and the last part of (iii) completes the proof of (iv). ∎

Clearly we have given a direct construction of the canonical measure, $R_t(x_0, \cdot)$, of X_t under $\mathbb{P}_{\delta_{x_0}}$. In this case $M \equiv 0$. For general γ, g it is not hard to modify the above to recover the canonical representation from our convergence theorem. We leave this as Exercise II.7.1 below. In general M will not be 0 as can readily be seen by taking $\gamma \equiv 0$.

Exercise II.7.1. Let X be a (Y, γ, g)-DW superprocess starting at δ_{x_0} where $\gamma \in C_b(E)_+$ and $g \in C_b(E)$. Extend the proof of Theorem II.7.2 to show there is an $M_t(x_0, \cdot) \in M_F(E)$ and a σ-finite measure $R_t(x_0, \cdot)$ on $M_F(E) - \{0\}$ such that $\int \nu(1)R_t(x_0, d\nu) < \infty$ and

$$\mathbb{E}_{\delta_{x_0}}\left(\exp\left(-X_t(\phi)\right)\right) = \exp\left\{-M_t(x_0, \phi) - \int 1 - e^{-\nu(\phi)}R_t(x_0, d\nu)\right\} \quad \forall \phi \in (b\mathcal{E})_+.$$

Hint. Recalling (II.7.6), let $\varepsilon > 0$, $\phi \in C_b(E)_+$, and write

$$\begin{aligned} \int \Big(1 - e^{-X^N_t(\phi)}\Big)NdP^N_{\delta_{x_0}/N} &= N\int \Big(1 - e^{-X^N_t(\phi)} - X^N_t(\phi)\Big)\, 1(X^N_t(1) \le \varepsilon)dP^N_{\delta_{x_0}/N} \\ (*) \qquad &\quad + \int X^N_t(\phi)1\left(X^N_t(1) \le \varepsilon\right) NdP^N_{\delta_{x_0}/N} \\ &\quad + \int (1 - e^{-X^N_t(\phi)})1(X^N_t(1) > \varepsilon)NdP^N_{\delta_{x_0}/N}. \end{aligned}$$

Show that the first term goes to 0 as $\varepsilon \downarrow 0$ uniformly in N and that

$$\Big\{\int X^N_t(\cdot)1(X^N_t(1) \le \varepsilon)NdP^N_{\delta_{x_0}/N} : N \in \mathbb{N}\Big\}$$

and

$$\{NP^N_{\delta_{x_0}/N}\big(X^N_t \in \cdot, X^N_t(1) > \varepsilon\big) : N \in \mathbb{N}\}$$

are tight on E and $M_F(E)$, respectively. Now let $N \to \infty$ through an appropriate subsequence and then $\varepsilon = \varepsilon_k \downarrow 0$ in $(*)$ to obtain the desired decomposition.

Theorem II.7.2 and (II.7.7) imply

$$P^N_{\delta_{x_0}/N}\left(X^N_t \in \cdot \mid X^N_t \neq 0\right) \overset{w}{\Rightarrow} R_t(x_0,\cdot)/R_t(x_0,1), \tag{II.7.10}$$

that is, $R_t(x_0,\cdot)$, when normalized, is the law of a cluster at time t of descendants of a common ancestor at $t = 0$ conditioned on the existence of such descendants. Note that Yaglom's Theorem II.1.1(b) is immediate from (II.7.10) and Theorem II.7.2(iii).

Exercise II.7.2. (a) If X is a $(Y,\gamma,0)$-DW-superprocess under $\mathbb{P}_{X_0}$ with $\gamma > 0$ constant, prove that $\varepsilon^{-1}\mathbb{P}_{\varepsilon\delta_{x_0}}(X_t \in \cdot, X_t \neq 0) \overset{w}{\Rightarrow} R_t(x_0,\cdot)$ as $\varepsilon \downarrow 0$ on $M_F(E)$.
Hint. Use Theorem II.7.2 (ii) to show convergence of the corresponding Laplace functionals.

(b) If T_t is the semigroup of X, show that $R_t(x_0,\psi) = \int T_{t-\tau}\psi(\nu)R_\tau(x_0,d\nu)$ for all $0 < \tau \leq t$, $x_0 \in E$, and ψ bounded measurable on $M_F(E) - \{0\}$.
Hint. Use (a), first consider $\psi \in C_b(M_F(E))$ such that $\psi(0) = 0$, and recall that $T_t : C_b(M_F(E)) \to C_b(M_F(E))$.

If $X_0 \in M_F(E)$, let Ξ^{t,X_0} be a Poisson point process on $M_F(E) - \{0\}$ with intensity $\int R_t(x_0,\cdot)dX_0(x_0)$. Theorem II.7.2(ii) implies that

$$\int \nu\ \Xi^{t,X_0}(d\nu) \quad \text{is equal in law to} \quad \mathbb{P}_{X_0}(X_t \in \cdot). \tag{II.7.11}$$

In view of (II.7.7) and (II.7.10) we see that this representation decomposes X_t according to the Poisson number of ancestors at time 0 with descendants alive at time t. This perspective will allow us to extend Theorem II.7.2 and this Poisson decomposition to the sample paths of X. Indeed, (II.7.1) shows that infinite divisibility is valid on the level of sample paths.

Let $\zeta : \Omega_D \to [0,\infty]$ be given by $\zeta(X) = \inf\{t > 0 : X_t = 0\}$ and define

$$\Omega^{Ex} = \{X \in \Omega_D : X_0 = 0, \zeta > 0, X_t \equiv 0 \quad \forall t \geq \zeta\},$$
$$\Omega^{Ex}_C = \left\{X \in \Omega^{Ex} : X_\cdot \quad \text{is continuous}\right\} \subset \Omega_X,$$

equipped with the subspace topologies they inherit from Ω_D and Ω_X, respectively. If $\{\mathbb{N}_k : k \in \mathbb{N}\cup\{\infty\}\}$ are measures on Ω^{Ex} we write $\mathbb{N}_k \overset{w}{\Rightarrow} \mathbb{N}_\infty$ on Ω^{Ex} if $\mathbb{N}_k(\zeta > t) < \infty$ for all $k \in \mathbb{N}\cup\{\infty\}$ and $t > 0$, and

$$\mathbb{N}_k(X \in \cdot, \zeta > t) \overset{w}{\Rightarrow} \mathbb{N}_\infty(X \in \cdot, \zeta > t) \text{ as } k \to \infty, \text{ as finite measures on } \Omega_D\ \forall t > 0.$$

Theorem II.7.3. (a) For each $x_0 \in E$ there is a σ-finite measure, $\mathbb{N}_{x_0}$, on Ω^{Ex}_C such that $NP_{\delta_{x_0}/N}(X^N_\cdot \in \cdot) \overset{w}{\Rightarrow} \mathbb{N}_{x_0}$ on Ω^{Ex}.
(b) For all $t > 0$, $\mathbb{N}_{x_0}(X_t \in \cdot, \zeta > t) = R_t(x_0,\cdot)$.
(c) Let Ξ be a Poisson point process on Ω^{Ex}_C with intensity $\mathbb{N}_{x_0}$. Then $X_t = \int \nu_t d\Xi(\nu)$, $t > 0$, has the law of a $(Y,\gamma,0)$-DW-superprocess starting at δ_{x_0}.

Remark II.7.4. Note that (II.7.7) and the equality $\mathbb{N}_{x_0}(\zeta > t) = R_t(x_0,1) = 2/\gamma t$ allow us to use (a) to conclude

$$P_{\delta_{x_0}/N}\left(X^N \in \cdot \mid X^N_t \neq 0\right) \overset{w}{\Rightarrow} \mathbb{N}_{x_0}\left(X \in \cdot \mid \zeta > t\right) \quad \text{on} \quad \Omega_D \quad \forall t > 0.$$

In this way $\mathbb{N}_{x_0}$ describes the time evolution of a cluster starting from a single ancestor at x_0 given that it survives for some positive length of time. We call $\mathbb{N}_{x_0}$ the canonical measure of the process X. It has been studied by El Karoui and Roelly (1991) and Li and Shiga (1995). A particularly elegant construction of $\mathbb{N}_{x_0}$ in terms of Le Gall's snake may be found in Chapter IV of Le Gall (1999). The reader may want to skip the proof of Theorem II.7.3 on a first reading.

Proof. Let $X^N_\cdot$ be as before under $P^N_{\delta_{x_0}}$, and for $i \le N$ let $X^{N,i}_t = \frac{1}{N} \sum_{\alpha \sim t, \alpha_0 = i} \delta_{Y^\alpha_t}$ be the portion of X^N_t descending from the i^{th} initial ancestor. Fix $t > 0$ and set $\Lambda^N_t = \{i \le N : X^{N,i}_t \ne 0\}$. The mutual independence of $\mathcal{G}_i = \sigma(Y^\alpha, N^\alpha : \alpha > i)$, $i = 1 \dots, N$, shows that conditional on Λ^N_t, $\{X^{N,i}_\cdot : i \in \Lambda^N_t\}$ are iid with law $P^N_{\delta_{x_0}/N}(X^N_\cdot \in \cdot \mid X^N_t \ne 0)$. We also have

$$(II.7.12) \qquad X^N_{t+\cdot} = \sum_{i \in \Lambda^N_t} X^{N,i}_{t+\cdot}.$$

Clearly $|\Lambda^N_t| = \text{card}(\Lambda^N_t)$ is binomial $(N, P_{\delta_{x_0}/N}(X^N_t \ne 0))$ and so by (II.7.7), converges weakly to a Poisson random variable Λ_t with mean $2/\gamma t$. The left side of (II.7.12) converges weakly on Ω_D to $\mathbb{P}_{\delta_{x_0}}(X_{t+\cdot} \in \cdot)$ (use the fact that the limit is a.s. continuous) and so for $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset \Omega_D$ such that $P^N_{\delta_{x_0}}(X^N_{t+\cdot} \in K^c_\varepsilon) < \varepsilon$ for all N. Use (II.7.12) to see that this means that for all N

$$\varepsilon > P^N_{\delta_{x_0}}\left(X^N_{t+\cdot} \in K^c_\varepsilon, |\Lambda^N_t| = 1\right) = P^N_{\delta_{x_0}/N}\left(X^N_{t+\cdot} \in K^c_\varepsilon | X^N_t \ne 0\right) P(|\Lambda^N_t| = 1).$$

This proves tightness of $\{P^N_{\delta_{x_0}/N}(X^N_{t+\cdot} \in \cdot \mid X^N_t \ne 0) : N \in \mathbb{N}\}$ on Ω_D because $\lim_{N\to\infty} P\big(|\Lambda^N_t| = 1\big) = (2/\gamma t)e^{-2/\gamma t} > 0$. Let P^t be any limit point on Ω_D. If Λ_t is the above Poisson random variable and conditional on Λ_t, $\{X^i_\cdot : i \le \Lambda_t\}$ are iid with law P^t, then we may let $N \to \infty$ through an appropriate subsequence in (II.7.12) and conclude

$$(II.7.13) \qquad \sum_{i=1}^{\Lambda_t} X^i_\cdot \quad \text{has law} \quad \mathbb{P}_{\delta_{x_0}}(X_{t+\cdot} \in \cdot).$$

Note that $P(\Lambda_t \ge 1) = 1 - e^{-2/\gamma t} = \mathbb{P}_{\delta_{x_0}}(X_t \ne 0)$ (recall (II.5.12)). From this and the above we may conclude that $\sum_{i=1}^{\Lambda_t} X^i_\cdot = 0$ iff $\Lambda_t = 0$ and therefore

$$P\left(\Lambda_t \ge 1, \sum_{i=1}^{\Lambda_t} X^i_\cdot \in \cdot\right) = \mathbb{P}_{\delta_{x_0}}\left(X_{t+\cdot} \in \cdot, X_t \ne 0\right).$$

The measure on the right is supported on

$$\Omega'_X = \{X \in \Omega_X : X_0 \ne 0, \zeta > 0, X_s = 0 \text{ all } s \ge \zeta\},$$

and so the same must be true of $P^t = \mathcal{L}(X^i_\cdot)$ as $P(\Lambda_t = 1, X^1_0 \notin \Omega'_X) = 0$ by the above.

If $0 \le t_1 < \ldots < t_k$ and $\phi_i \in C_b(E)_+$ for $1 \le i \le k$, (II.7.13) shows that

$$\mathbb{P}_{\delta_{x_0}}\Big(\exp\Big\{-\sum_1^k X_{t_j+t}(\phi_j)\Big\}\Big) = \exp\Big\{-\int 1-\exp\Big(-\sum_1^k \nu_{tj}(\phi_j)\Big)dP^t(\nu)2/\gamma t\Big\}.$$

This uniquely determines $\int \exp\big(-\sum\limits_1^k \nu_{tj}(\phi_j)\big)dP^t(\nu)$ and hence the finite-dimensional distributions of P^t by a now familiar monotone class argument. We have shown (use (II.7.7))

$$(II.7.14)\qquad \begin{aligned} &NP^N_{\delta_{x_0}/N}\left(X^N_{t+\cdot} \in \cdot, X^N_t \ne 0\right)\\ &= NP^N_{\delta_{x_0}/N}\left(X^N_t \ne 0\right) P^N_{\delta_{x_0}/N}\left(X^N_{t+\cdot} \in \cdot \mid X^N_t \ne 0\right)\\ &\stackrel{w}{\Rightarrow} \frac{2}{\gamma t}P^t \qquad \text{on} \quad \Omega_D, \end{aligned}$$

where the limit is supported on Ω'_X.

To handle the small values of t we need a definition and a Lemma.

Notation. $\mathrm{Lip}_1 = \{\phi : E \to \mathbb{R} : \|\phi\| \le 1,\ |\phi(x) - \phi(y)| \le \rho(x,y)\}$ (ρ is a fixed complete metric on E).

Definition. The Vasershtein metric $d = d_\rho$ on $M_F(E)$ is

$$d(\mu,\nu) = \sup\left\{|\mu(\phi) - \nu(\phi)| : \phi \in \mathrm{Lip}_1\right\}.$$

Then d is a complete metric which induces the weak topology on $M_F(E)$ (e.g. see Ethier and Kurtz (1986), p. 150, Exercise 2). It only plays an incidental role here but will be important in treating models with spatial interactions in Chapter V.

Redefine $X^N_\cdot$ near $t = 0$ by

$$\tilde X^N_t = \begin{cases} X^N_t & \text{if } t \ge N^{-3} \\ tN^3 X^N_{N^{-3}} & \text{if } t \in [0, N^{-3}) \end{cases}.$$

Lemma II.7.5. (a) $NP^N_{\delta_{x_0}/N}\big(\sup\limits_{s\le\delta} X^N_s(1) > \varepsilon\big) \le 4\gamma\delta\varepsilon^{-2}$ for all $\delta, \varepsilon > 0$ and $N \ge 2/\varepsilon$.

(b) There are $N_0 \in \mathbb{N}$ and $\delta_0 > 0$ such that

$$NP^N_{\delta_{x_0}/N}\left(\sup_{s\le\delta} \tilde X^N_s(1) > \delta^{1/5}\right) \le \delta^{1/2} \quad \text{for} \quad 0 < \delta \le \delta_0 \quad \text{and} \quad N \ge N_0.$$

(c) $NP^N_{\delta_{x_0}/N}\left(\sup\limits_t d\left(X^N_t, \tilde X^N_t\right) \ge 4/N\right) \le \gamma N^{-1}$.

Proof. (a) Use $(MP)^N$ in Section II.4 and (II.4.5) with $\phi = 1$, $g_N = 0$ and γ_N constant to see that under $P^N_{\delta_{x_0}/N}$, $X^N_t(1)$ is a martingale with predictable square function $\gamma\int\limits_0^t X^N_s(1)ds$. The weak L^1 inequality and Lemma II.3.3 (a) imply that for

$N > 2/\varepsilon$

$$NP^N_{\delta_{x_0}/N}\left(\sup_{s\le\delta} X^N_s(1) > \varepsilon\right) \le NP^N_{\delta_{x_0}/N}\left(\sup_{s\le\delta} X^N_s(1) - X^N_0(1) \ge \varepsilon/2\right)$$
$$\le N4\varepsilon^{-2}P^N_{\delta_{x_0}/N}\left(\gamma\int_0^\delta X^N_s(1)ds\right) = 4\gamma\delta\varepsilon^{-2}.$$

(b) Assume first $\delta \ge N^{-3}$. Then for $N \ge N_0$, $2\delta^{-1/5} \le N$ and so by (a) for $\delta < \delta_0$,

$$NP^N_{\delta_{x_0}/N}\left(\sup_{s\le\delta} \tilde{X}^N_s(1) > \delta^{1/5}\right) = NP^N_{\delta_{x_0}/N}\left(\sup_{N^{-3}\le s\le\delta} X^N_s(1) > \delta^{1/5}\right)$$
$$\le 4\gamma\delta^{3/5} \le \delta^{1/2}.$$

Assume now $\delta < N^{-3}$. Then the above probability equals

$$NP^N_{\delta_{x_0}/N}\left(\delta N^3 X^N_{N^{-3}}(1) > \delta^{1/5}\right) = NP^N_{\delta_{x_0}/N}\left(X^N_{N^{-3}}(1) > \delta^{-4/5}N^{-3}\right).$$

Our assumption on δ implies $2\delta^{4/5}N^3 < 2N^{3/5} < N$ for $N \ge N_0$ and so (a) bounds the above by

$$4\gamma N^{-3}N^6\delta^{8/5} < 4\gamma\delta^{3/5} < \delta^{1/2} \quad \text{for} \quad \delta \le \delta_0.$$

(c) If $f \in \mathrm{Lip}_1$ and $t < N^{-3}$,

$$\left|X^N_t(f) - \tilde{X}^N_t(f)\right| \le \|f\|\left[X^N_t(1) + tN^3X^N_{N^{-3}}(1)\right] \le 2\sup_{t\le N^{-3}} X^N_t(1).$$

This shows that the right-hand side is an upper bound for $\sup_t d(X^N_t, \tilde{X}^N_t)$ and an application of (a) completes the proof. ∎

We now complete the

Proof of Theorem II.7.3. Lemma II.7.5 (c) and (II.7.14) show that if $t_n \downarrow 0$ ($t_n > 0$) is fixed and $\varepsilon > 0$, we may choose $\tilde{K}^\varepsilon_n$ compact in Ω_D such that

$$(II.7.15) \qquad \sup_N NP^N_{\delta_{x_0}/N}\left(\tilde{X}^N_{t_n+\cdot} \notin \tilde{K}^\varepsilon_n, X^N_{t_n} \ne 0\right) < \varepsilon 2^{-n}.$$

For $t > 0$ define

$$K^\varepsilon_t = \left\{X \in \Omega_D : X_{t_n+\cdot} \in \tilde{K}^\varepsilon_n \ \ \forall t_n \le t \text{ and } \sup_{s\le 2^{-2n}} X_s(1) \le 2^{-2n/5} \text{ for all } n \ge 1/\varepsilon\right\}.$$

Lemma II.7.5(b) and (II.7.15) show that for $\varepsilon < \varepsilon_0$ and $N \geq N_0$

$$\begin{aligned} &NP^N_{\delta_{x_0}/N}\left(\tilde{X}^N \notin K^\varepsilon_t, X^N_t \neq 0\right) \\ &\leq \sum_n 1(t_n \leq t) NP^N_{\delta_{x_0}/N}\left(\tilde{X}^N_{t_n+\cdot} \notin \tilde{K}^\varepsilon_n, X^N_{t_n} \neq 0\right) \\ &\qquad\qquad + \sum_{n\geq 1/\varepsilon} NP^N_{\delta_{x_0}/N}\big(\sup_{s\leq 2^{-2n}} \tilde{X}^N_s(1) > 2^{-2n/5}\big) \\ &\leq \sum_{n=1}^\infty \varepsilon 2^{-n} + \sum_{n\geq 1/\varepsilon} 2^{-n} \leq 2\varepsilon. \end{aligned}$$

It is a routine Skorokhod space exercise to check that $\overline{K^\varepsilon_t}$ is a compact subset of $\Omega^0_D = \{X \in \Omega_D : X_0 = 0\}$. This together with Lemma II.7.5(c) shows that $\{NP^N_{\delta_{x_0}/N}(X^N \in \cdot, \zeta > t) : N \in \mathbb{N}\}$ is relatively compact in Ω_D and all limit points are supported on Ω^0_D.

Fix $t_0 > 0$. Choose $N_k \to \infty$ such that

$$(II.7.16) \qquad NP^N_{\delta_{x_0}/N}\left(X^N \in \cdot, \zeta > t_0\right) \stackrel{w}{\Rightarrow} \mathbb{N}^{t_0} \text{ on } \Omega_D \text{ as } N \to \infty \text{ through } \{N_k\}.$$

To ease our notation we write N for N_k and $\mathbb{Q}_N$ for $NP^N_{\delta_{x_0}/N}$. By taking a further subsequence we may assume (II.7.16) holds with t_m in place of t_0 (recall $t_m \downarrow 0$). Clearly $\mathbb{N}^{t_m}(\cdot)$ are increasing in m and so we may define a measure on Ω_D by $\mathbb{N}_{x_0}(A) = \lim_{m\to\infty} \mathbb{N}^{t_m}(A)$. Let $t_m < t$. Theorem II.7.2 implies that

$$\lim_{\varepsilon\downarrow 0} \overline{\lim_{N\to\infty}}\, \mathbb{Q}_N\left(X^N_t(1) \in (0,\varepsilon), \zeta > t_m\right) \leq \lim_{\varepsilon\downarrow 0} R_t\left(x_0, \{\nu : \nu(1) \in (0,\varepsilon)\}\right) = 0.$$

A standard weak convergence argument now gives

$$\begin{aligned} \lim_{N\to\infty} \mathbb{Q}_N\left(X^N \in \cdot, \zeta > t\right) &= \lim_{N\to\infty} \mathbb{Q}_N\left(X^N \in \cdot, X^N_t(1) > 0, \zeta > t_m\right) \\ &= \mathbb{N}^{t_m}\left(\cdot\ , X_t(1) > 0\right). \end{aligned}$$

This shows that the measure on the right side is independent of m and so the above implies

$$(II.7.17) \qquad \mathbb{Q}_N\left(X^N \in \cdot, \zeta > t\right) \stackrel{w}{\Rightarrow} \mathbb{N}_{x_0}\left(\cdot\ , X_t(1) > 0\right) \qquad \forall t > 0,$$

and in particular (take $t = t_m$)

$$(II.7.18) \qquad \mathbb{N}_{x_0}\left(\cdot, X_{t_m}(1) > 0\right) = \mathbb{N}^{t_m}(\cdot\ ), \quad m = 0, 1, 2, \ldots.$$

(II.7.17) shows that the measures $\mathbb{N}_{x_0}\big(\cdot\ , X_t(1) > 0\big)$ are decreasing in t and this implies for each $s < t$, $\mathbb{N}_{x_0}$ a.s. $X_s = 0$ implies $X_t = 0$. Right-continuity implies

$$(II.7.19) \qquad X_s = 0 \Rightarrow X_t = 0 \quad \forall s < t \quad \mathbb{N}_{x_0}\text{-a.s.}$$

This implies $\{\zeta = 0\} \subset \bigcap_m \{X_{t_m} = 0\}$ $\mathbb{N}_{x_0}$-a.s., and therefore $\mathbb{N}^{t_m}$-a.s. Therefore

$$\mathbb{N}^{t_m}(\zeta = 0) \leq \mathbb{N}^{t_m}(X_{t_m} = 0) = 0,$$

the last by (II.7.18). It follows that $\mathbb{N}_{x_0}(\zeta = 0) = \lim_{m\to 0} \mathbb{N}^{t_m}(\zeta = 0) = 0$ which, together with (II.7.19) shows that $\mathbb{N}_{x_0}$ is supported by Ω^{Ex}. (II.7.17) may therefore be written as

$$(II.7.20) \qquad \mathbb{Q}_N\left(X^N \in \cdot, \zeta > t\right) \overset{w}{\Rightarrow} \mathbb{N}_{x_0}\left(\cdot, \zeta > t\right) \qquad \forall t > 0.$$

The convergence in (II.7.14) and the above together imply

$$(II.7.21) \qquad \mathbb{N}_{x_0}\left(X_{t+\cdot} \in \cdot, \zeta > t\right) = \frac{2}{\gamma t} P^t(\cdot)$$

(Note that (II.7.20) alone would not give this if t is a point of discontinuity, but as the limit in (II.7.14) exists we only need identify the finite-dimensional distributions in terms of $\mathbb{N}_{x_0}$ and (II.7.20) is enough for this.) This implies $\mathbb{N}_{x_0}$-a.s. continuity of $X_{t+\cdot}$ for all $t > 0$ (recall $P^t(\Omega'^c_X) = 0$) and hence shows $\mathbb{N}_{x_0}$ is supported on Ω_C^{Ex}. (II.7.21) also identifies the finite dimensional distributions of $\mathbb{N}_{x_0}$ and so by (II.7.18) with $m = 0$ we may conclude that all limit points in (II.7.16) equal $\mathbb{N}_{x_0}(\cdot, \zeta > t_0)$. This proves (a). (b) is then immediate from Theorem II.7.2(i).

Let Ξ be as in (c). Note that $\Xi\big(\{\nu : \nu_t \neq 0\}\big)$ is Poisson with mean $\mathbb{N}_{x_0}(\zeta > t) = 2/\gamma t$ (by (II.7.21)) and so $\int \nu_{t+\cdot}\Xi(d\nu) \overset{\mathcal{D}}{\equiv} \sum_{i=1}^{\Lambda_t} X^i_{t+\cdot}$, where Λ_t is Poisson $(2/\gamma t)$ and given Λ_t, $\{X^i_{t+\cdot} : i \le \Lambda_t\}$ are iid with law $\mathbb{N}_{x_0}(X_{t+\cdot} \mid \zeta > t) = P^t$ (by (II.7.21)). Compare this with (II.7.13) and let $t \downarrow 0$ to complete the proof of (c). ∎

8. Historical Processes.

We return to the path-valued setting of Example II.2.4(c) under the assumption

$$(PC) \qquad x \to P^x \quad \text{is continuous.}$$

In addition to the $\hat{E}$-valued BSMP W_t with laws $(\hat{P}_{\tau,y})$ described there, we introduce probabilities $\left\{P_{\tau,y} : (\tau, y) \in \hat{E}\right\}$ on $D(E)$ by

$$P_{\tau,y}(A) = P^{y(\tau)}\left(\{w : (y/\tau/w) \in A\}\right).$$

If $W_\cdot$ has law $\hat{P}_{\tau,y}$ and Y has law $P_{\tau,y}$ then

$$(II.8.1) \qquad (W_t)_{t\ge 0} \overset{\mathcal{D}}{\equiv} \left(\tau + t, Y^{\tau+t}\right)_{t\ge 0}.$$

Let $\hat{g} \in C_b(\hat{E})$, $\hat{\gamma} \in C_b(\hat{E})_+$, and for $\tau \ge 0$ define

$$M_F^\tau(D) = \left\{m \in M_F\big(D(E)\big) : y^\tau = y \quad m - a.a.\ y\right\}.$$

Fix $m \in M_F^\tau(D)$ so that $\delta_\tau \times m \in M_F(\hat{E})$ and let $\hat{X}$ be the $(W, \hat{\gamma}, \hat{g})$-DW superprocess starting at $\delta_\tau \times m$ with law $\hat{\mathbb{P}}_{\tau,m} (\equiv \hat{\mathbb{P}}_{\delta_\tau\times m})$ on the canonical space of continuous $M_F(\hat{E})$-valued paths. Introduce

$$\Omega_H[\tau, \infty) = \left\{H_\cdot \in C\Big([\tau, \infty), M_F\big(D(E)\big)\Big) : H_t \in M_F^t(D) \quad \forall t \ge \tau\right\},$$

and let $\Omega_H = \Omega_H[0,\infty)$ with its Borel σ-field $\mathcal{F}_H$. Let $\Pi : \hat{E} \to D(E)$ be the projection map and define an $M_F\big(D(E)\big)$-valued process $(H_t, t \geq \tau)$ by

$$H_{\tau+t}(A) = \hat{X}_t\big(\Pi^{-1}(A)\big).$$

Lemma II.8.1. $\hat{X}_t = \delta_{\tau+t} \times H_{\tau+t}$ $\forall t \geq 0$ and $H_t \in M_F^t(D)$ $\forall t \geq \tau$ $\hat{\mathbb{P}}_{\tau,m}$-a.s.

Proof. Let $\hat{P}_t^{\hat{g}} f(\tau, y) = \hat{P}_{\tau,y}\big(\exp\{\int_0^t \hat{g}(W_s)\,ds\} f(W_t)\big)$ and

$$\Lambda(t) = \{(u, y) \in \hat{E} : u \neq \tau + t\}.$$

Then by Exercise II.5.2(b)

$$\mathbb{P}_{\tau,m}\left(\hat{X}_t\big(\Lambda(t)\big)\right) = \int \hat{P}_t^{\hat{g}}(1_{\Lambda(t)})(\tau, y)\,dm(y) = 0$$

because $W_t = \big(\tau + t, Y^{\tau+t}\big)$ $\hat{P}_{\tau,y}$-a.s. This shows $\hat{X}_t = \delta_{\tau+t} \times H_{\tau+t}$ $\mathbb{P}_{\tau,m}$-a.s. for each $t \geq 0$ and hence for all $t \geq 0$ a.s. by the right-continuity of both sides. Since $\hat{X}_t \in M_F(\hat{E})$ $\forall t \geq 0$ a.s., the second assertion follows immediately. ∎

The process of interest is the $M_F(D)$-valued process H. We abuse our notation and also use H_t to denote the coordinate variables on Ω_H and let

$$\mathcal{F}^H[s, t+] = \bigcap_{n=1}^{\infty} \sigma(H_r : s \leq r \leq t + 1/n).$$

Define $\mathbb{Q}_{\tau,m}$ on $\big(\Omega_H, \mathcal{F}^H[\tau,\infty)\big)$ by $\mathbb{Q}_{\tau,m}(\cdot) = \hat{\mathbb{P}}_{\tau,m}(H \in \cdot)$, where H is as in Lemma II.8.1. The fact that $\hat{X}$ is a BSMP easily shows that

$$H \equiv (\Omega_H, \mathcal{F}_H, \mathcal{F}^H[\tau, t+], H_t, \mathbb{Q}_{\tau,m})$$

is an inhomogeneous Borel strong Markov process (IBSMP) with continuous paths in $M_F^t(D) \subset M_F(D)$. This means

(i) $\forall u > 0$ and $A \in \mathcal{F}^H[u,\infty)$ $(\tau, m) \to \mathbb{Q}_{\tau,m}(A)$ is Borel measurable on $\{(\tau, m) : m \in M_F^\tau(D), \tau \leq u\}$.

(ii) $H_\tau = m$, $H_t \in M_F^t(D)$ $\forall t \geq \tau$, and $H_\cdot$ is continuous $\mathbb{Q}_{\tau,m}$-a.s.

(iii) If $m \in M_F^\tau(D)$, $\psi \in b\mathcal{B}\big([\tau,\infty] \times M_F(D)\big)$ and $T \geq \tau$ is an $(\mathcal{F}^H[\tau, t+])_{t \geq \tau}$-stopping time, then

$$\mathbb{Q}_{\tau,m}\left(\psi(T, H_{T+\cdot}) \mid \mathcal{F}^H[\tau, T+]\right)(\omega) = \mathbb{Q}_{T(\omega), H_T(\omega)}\Big(\psi\big(T(\omega), H_{T(\omega)+\cdot}\big)\Big)$$
$$\mathbb{Q}_{\tau,m} - \text{a.s. on} \quad \{T < \infty\}.$$

This is a simple consequence of the fact that $\hat{X}$ is a BSMP (only (iii) requires a bit of work) and the routine proof is contained in Dawson-Perkins (1991) (Proof of Theorem 2.1.5 in Appendix 1). We call H the $(Y, \hat{\gamma}, \hat{g})$-historical process.

Of course it is now a simple matter to interpret the weak convergence theorem (Theorem II.5.1(c)), local martingale problem $(LMP)_\nu$, and Laplace equation (LE) for $\hat{X}$, in terms of H.

To link the weak convergence result with that for the (Y, γ, g)-superprocess consider the special case where $\hat{\gamma}(t, y) = \gamma\big(y(t)\big)$, $\hat{g}(t, y) = g\big(y(t)\big)$ for some $g \in$

$C_b(E)$ and $\gamma \in C_b(E)_+$, $\tau = 0$ and $m \in M_F^0(D)$. Note that we can, and shall, consider m as a finite measure on E. Note also that $\hat{\gamma}$ and $\hat{g}$ are continuous on $\hat{E}$ (but not necessarily on $\mathbb{R}_+ \times D$) – see Exercise II.8.1 below. It is then natural to assume our approximating branching mechanisms $\hat{\nu}^N\big((t,y),\cdot\big) = \nu^N\big(y(t),\cdot\big)$ where $\{\nu^N\}$ satisfy (II.3.1). Let $\{Y^\alpha : \alpha \sim t\}$ be the system of branching Y-processes constructed in Section II.3. If $W_t^\alpha = (t, Y^\alpha_{\cdot\wedge t})$, then $\{W^\alpha : \alpha \sim t\}$ is the analogous system of branching W-processes and so Theorem II.5.1(c) implies

$$(II.8.2) \qquad \hat{X}_t^N(\cdot) = \frac{1}{N}\sum_{\alpha\sim t}\delta_{W_t^\alpha} \overset{w}{\Rightarrow} \hat{X}_t \quad \text{on} \quad D\big(\mathbb{R}_+, M_F(\hat{E})\big),$$

where $\hat{X}$ has law $\mathbb{P}_{0,m}$.

Recall $H_t^N \in M_F\big(D(E)\big)$ is defined by $H_t^N = \hat{X}_t^N\big(\Pi^{-1}(\cdot)\big)$, i.e.,

$$H_t^N = \frac{1}{N}\sum_{\alpha\sim t}\delta_{Y^\alpha_{\cdot\wedge t}},$$

and so, taking projections in (II.8.2), we have

$$(II.8.3) \qquad \mathbb{P}(H^N_\cdot \in \cdot) \overset{w}{\Rightarrow} \mathbb{Q}_{0,m}(\cdot).$$

Remark II.8.2. It is possible to prove this weak convergence result without the continuity assumption (PC) and to prove Theorem II.5.1(c) with the assumption (II.2.2) $P_t : C_b \to C_b$ replaced by the weaker condition (QLC) (i.e. Y is a Hunt process). For γ constant and $g \equiv 0$ these results are proved in Theorems 7.15 and 7.13, respectively, of Dawson-Perkins (1991). Our proof of the compact containment condition (i) can be used to simplify the argument given there. Without our continuity assumptions one must work with the fine topology and use a version of Lusin's theorem. The processes of interest to us satisfy our continuity conditions and so we will not discuss these extensions.

It is a relatively simple matter to take projections in (II.8.2) (or (II.8.3)) and compare with Theorem II.5.1(c) to see that
$(II.8.4)$

$X_t \equiv H_t(y_t \in \cdot)$ is a (Y,γ,g)-DW superprocess starting at m under $\mathbb{Q}_{0,m}$.

We leave this as Exercise II.8.1. See Exercise II.8.3 for another approach.

Exercise II.8.1. (a) Define $\hat{\Pi} : \mathbb{R}_+ \times D(E) \to E$ by $\hat{\Pi}(t,y) = y(t)$. Show that $\hat{\Pi}$ is not continuous but its restriction to $\hat{E}$ is.
Hint. On $\hat{E}$, $\hat{\Pi}(t,y) = y(T)$ for all $T \geq t$.

(b) For $H \in \Omega_H$ define $\tilde{\Pi}(H)(t) = H_t(y_t \in \cdot) \in M_F(E)$. Show that $\tilde{\Pi} : \Omega_H \to \Omega_X$ and is continuous.
Hints. (i) Show

$$\begin{aligned} T :&\Omega_H \to C(\mathbb{R}_+, M_F(\hat{E})) \\ &H_t \to \delta_t \times H_t \end{aligned}$$

is continuous.
(ii) Show that $\tilde{\Pi}(H)_t = T(H)_t \circ \hat{\Pi}^{-1}$.

(c) Use either (II.8.2) and (a), or (II.8.3) and (b), to prove that under $\mathbb{Q}_{0,m}$, $\tilde{\Pi}(H_\cdot)$ is a (Y,γ,g)-DW superprocess starting at $m \in M_F(E)$.

Consider now the version of (MP) which characterizes $\mathbb{Q}_{\tau,m}$. For $s \le t$ and ϕ in $b\mathcal{D}\Big(\mathcal{D} = \mathcal{B}\big(D(E)\big)\Big)$ let $P_{s,t}\phi(y) = P_{s,y}\big(\phi(Y^t)\big)$ be the inhomogeneous semigroup associated with the path-valued process Y^t. If $\hat{A}$ is the weak generator of W and $\hat{A}\phi(s,y) \equiv \hat{A}_s\phi(y)$, then it is easy to see from Proposition II.2.1 that for $\phi \in C_b(\hat{E})$,

$$\phi \in \mathcal{D}(\hat{A}) \Leftrightarrow (P_{s,s+h}\phi_{s+h} - \phi_s)\,/h \overset{bp}{\to} \hat{A}_s\phi \text{ as } h \downarrow 0 \text{ for some } \hat{A}_s\phi(y) \text{ in } C_b(\hat{E})$$

$$\Leftrightarrow \text{ For some } \hat{A}_s\phi(y) \text{ in } C_b(\hat{E}),\ \phi(t,Y^t) - \phi(s,Y^s) - \int_s^t \hat{A}_r\phi(Y^r)\,dr,$$

$$t \ge s, \text{ is a } P_{s,y}\text{-martingale } \forall (s,y) \in \hat{E}.$$

If $m \in M_F^\tau(D)$, an $(\mathcal{F}_t)_{t\ge\tau}$-adapted process $(H_t)_{t\ge\tau}$ with sample paths in $\Omega_H[\tau,\infty)$ defined on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge\tau},\mathbb{P})$ satisfies $(HMP)_{\tau,m}$ iff $H_\tau = m$ a.s. and

$$\forall \phi \in \mathcal{D}(\hat{A}) \quad M_t(\phi) = H_t(\phi_t) - H_\tau(\phi_\tau) - \int_\tau^t H_s(\hat{A}_s\phi + \hat{g}_s\phi_s)\,ds \quad \text{is a}$$

$$\text{continuous } (\mathcal{F}_t)\text{-martingale with} \quad \langle M(\phi)\rangle_t = \int_\tau^t H_s(\hat{\gamma}_s\phi_s^2)\,ds \quad \forall t \ge \tau \quad \text{a.s.}$$

The following is immediate from Theorem II.5.1, applied to $\hat{X}$. As we are considering deterministic initial conditions we may work with a martingale problem rather than a local martingale problem (recall Remark II.5.5(2)).

Theorem II.8.3. (a) $(HMP)_{\tau,m}$ is well-posed. $\mathbb{Q}_{\tau,m}$ is the law of any solution to $(HMP)_{\tau,m}$.

(b) If K satisfies $(HMP)_{\tau,m}$ on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge\tau}\mathbb{P})$ and $T \ge \tau$ is an $(\mathcal{F}_t)_{t\ge\tau}$-stopping time, then

$$\mathbb{P}(K_{T+\cdot} \in A \mid \mathcal{F}_T)(\omega) = \mathbb{Q}_{T(\omega),K_T(\omega)}(H_{T(\omega)+\cdot} \in A) \quad \text{a.s.} \quad \forall A \in \mathcal{F}_H. \quad \blacksquare$$

We call a solution $(K_t, t \ge \tau)$ on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge\tau},\mathbb{P})$, an $(\mathcal{F}_t)$-historical process (or $(\mathcal{F}_t) - (Y,\hat{\gamma},\hat{g})$-historical process) starting at (τ,m).

The Feyman-Kac semigroup associated with Y^t and $\hat{g}$ is

$$P_{s,t}^{\hat{g}}\phi(y) = P_{s,y}\Big(\exp\Big\{\int_s^t \hat{g}(u,Y^u)\,du\Big\}\phi(Y^t)\Big) \qquad 0 \le s \le t, \quad \phi \in b\mathcal{D}.$$

The mean measure formula for DW-superprocesses (Exercise II.5.2 (b)) gives

$$\hat{\mathbb{P}}_{\tau,m}\left(\hat{X}_t(\phi)\right) = \int \hat{P}_{\tau,y}\Big(\exp\Big\{\int_0^t \hat{g}(W_s)\,ds\Big\}\phi(W_t)\Big)\,dm(y),$$

which by (II.8.1) and Lemma II.8.1 gives (set $\phi(t,y)=\psi(y^t)$ for $\psi\in b\mathcal{D}$)

$$(II.8.5)\qquad \mathbb{Q}_{\tau,m}\left(H_t(\psi)\right)=\int P^{\hat g}_{\tau,t}\psi(y)\,dm(y)\qquad t\ge\tau,\quad \psi\in b\mathcal{D}.$$

Let $\{\hat P_t\}$ denote the semigroup of W and let $\hat\phi$, $\hat f\in b\hat{\mathcal{E}}_+$ ($\hat{\mathcal{E}}$ is the Borel σ-field of $\hat E$). The Laplace equation for $\hat X$ (Theorem II.5.11) shows that if $\hat V_t$ is the unique solution of

$$(\widehat{ME})_{\hat\phi,\hat f}\qquad \hat V_t=\hat P_t\hat\phi+\int_0^t\hat P_s\left(\hat f+\hat g\hat V_{t-s}-\frac{\hat\gamma(\hat V_{t-s})^2}{2}\right)ds,$$

then

$$(\widehat{LE})\qquad \hat{\mathbb{P}}_{\tau,m}\Big(\exp\Big(-\hat X_t(\hat\phi)-\int_0^t\hat X_s(\hat f)\,ds\Big)\Big)=\exp\left\{-\int\hat V_t(\tau,y)\,dm(y)\right\}.$$

Let $D^s=\{y\in D(E):y=y^s\}$. Defining $V_{s,t}(y)=\hat V_{t-s}(s,y)$ $(s\le t, y\in D^s)$ establishes a one-to-one correspondence between solutions of $(\widehat{ME})_{\hat\phi,\hat f}$ and solutions of

$$(ME)_{\hat\phi,\hat f}\qquad V_{\tau,t}(y)=P_{\tau,t}(\hat\phi_t)(y)+\int_\tau^t P_{\tau,s}\left(\hat f_s+\hat g_sV_{s,t}-\hat\gamma_sV^2_{s,t}/2\right)(y)\,ds.$$

Note that in $(ME)_{\hat\phi,\hat f}$ we may fix t (and $\hat\phi_t$) and obtain an equation in $y\in D^\tau$, $\tau\le t$. Using Lemma II.8.1, and setting $\hat\phi(t,y)=\phi(y^t)$ for $\phi\in b\mathcal{D}_+$, and $\hat f(t,y)=f(t,y^t)$ for $f\in b(\mathcal{B}(\mathbb{R}_+)\times\mathcal{D})_+$, we readily translate $(\widehat{LE})$ into

Theorem II.8.4. Assume $\phi(y)$ and $f(t,y)$ are non-negative, bounded, Borel functions on $D(E)$ and $\mathbb{R}_+\times D(E)$, respectively. Let $V_{\tau,t}(y)$ be the unique solution of $(ME)_{\hat\phi,\hat f}$ with $\hat\phi$ and $\hat f$ as above. Then

$$(HLE)\qquad \mathbb{Q}_{\tau,m}\Big(\exp\Big\{-H_t(\phi)-\int_\tau^t H_s(f_s)\,ds\Big\}\Big)=\exp\left\{-\int V_{\tau,t}(y)\,dm(y)\right\}.$$

Exercise II.8.2. Assume $(H_t,t\ge\tau)$ is an $(\mathcal{F}_t)-(Y,\hat\gamma,0)$-historical process starting at (τ,m). Show that for any $\phi\in b\mathcal{D}$, $\int\phi(y^\tau)H_t(dy)$is a continuous $(\mathcal{F}_t)$-martingale.
Hint. The martingale property is easy (use (II.8.5)). For the continuity, start with ϕ continuous and then show the class of ϕ for which continuity holds is closed under $\xrightarrow{bp}$.

Exercise II.8.3. Let $m\in M_F^\tau(D)$, and assume $\hat g(s,y)=g(y(s))$ and $\hat\gamma(s,y)=\gamma(y(s))$ for some $g\in C_b(E)$ and $\gamma\in C_b(E)_+$. Let K satisfy $(HMP)_{\tau,m}$ on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge\tau},P)$. Define $X_t\in M_F(E)$ by $X_t(A)=K_{\tau+t}(\{y:y_{\tau+t}\in A\})$. From the hint in Exercise II.8.1(a) it is easy to see that $X.$ is a.s. continuous.

(a) If $\phi \in \mathcal{D}(A)$, prove that $\hat{\phi}(s,y) = \phi(y(s)),\ (s,y) \in \hat{E}$, defines a function in $\mathcal{D}(\hat{A})$ and $\hat{A}\phi(s,y) = A\phi(y(s))$.
(b) Show that X solves $(MP)^{g,\gamma,A}_{X_0}$ and conclude that $P(X \in \cdot) = \mathbb{P}_{X_0}$ is the law of the (Y,γ,g)-superprocess starting at $X_0 = m(y_\tau \in \cdot)$.

Since $\hat{X}_t$ is an infinitely divisible random measure, the same is true of H_t under $\mathbb{Q}_{\tau,m}$ for each $t \geq \tau$. We can therefore introduce the canonical measures from Section 7 in the historical setting. Assume $\hat{\gamma} \equiv \gamma$ is constant and $\hat{g} \equiv 0$. If $(\tau,y) \in \hat{E}$, let $\hat{R}_t(\tau,y)$ denote the canonical measure of $\hat{X}_t$ from Theorem II.7.2. Then Lemma II.8.1 and (II.7.11), applied to $\hat{X}$, imply that

$$\hat{R}_t(\tau,y)(\cdot) = \delta_{\tau+t} \times R_{\tau,\tau+t}(y,\cdot),$$

where (by Theorem II.7.2)

(a) $R_{\tau,t}(y,\cdot)$ is a finite measure on $M_F^t(D) - \{0\}$ which is Borel in $y \in D^\tau$, and satisfies $R_{\tau,t}(M_F^t(D)-\{0\}) = \dfrac{2}{\gamma(t-\tau)}$,

$$\int \psi(\nu(1))\nu(\phi)R_{\tau,t}(y,d\nu) = \int_0^\infty \psi(\gamma(t-\tau)z/2)ze^{-z}dzP_{\tau,y}(\phi(Y^t))$$

for any bounded Borel $\phi : D(E) \to \mathbb{R}$ and $\psi : \mathbb{R}_+ \to \mathbb{R}$, and $R_{\tau,t}(y,\{w : w^\tau \neq y\}) = 0$.

$$(II.8.6)\quad (b)\ \mathbb{Q}_{\tau,m}(\exp(-H_t(\phi))) = \exp\left(-\int\int 1 - e^{-\nu(\phi)}R_{\tau,t}(y,d\nu)m(dy)\right)$$

for any Borel $\phi : D(E) \to \mathbb{R}_+$.

(c) If $m \in M_F^\tau(D)$ and Ξ is a Poisson point process on $M_F(D) - \{0\}$ with intensity $\int R_{\tau,t}(y,\cdot)dm(y)$ then $\int \nu\Xi(d\nu)$ has law $\mathbb{Q}_{\tau,m}(H_t \in \cdot)$.

The fact that $w^\tau = y$ for $R_{\tau,t}(y)$-a.a. w (in (a)) is immediate from (c) and the corresponding property for H_t under $\mathbb{Q}_{\tau,\delta_y}$ (use (II.8.5) to see the latter).

The uniqueness of the canonical measure in Theorem II.7.1 and the fact that under $\mathbb{Q}_{\tau,\delta_y}$, $X_t(\cdot) = H_{\tau+t}(\{y' : y'_{\tau+t} \in \cdot\})$ is a $(Y,\gamma,0)$-superprocess starting at y_τ $(y \in D^\tau)$ by Exercise II.8.3, show that if $\hat{\Pi}_t(y) = y_t$ and $(R_t(x))_{x\in E}$ are the canonical measures for X, then

$$(II.8.7)\qquad R_{\tau,\tau+t}(y,\nu\circ\hat{\Pi}^{-1}_{\tau+t} \in \cdot) = R_t(y_\tau,\cdot).$$

III. Sample Path Properties of Superprocesses

1. Historical Paths and a Modulus of Continuity

Assume $\big(X,(\mathbb{P}_\mu)_{\mu\in M_F(E)}\big)$ is a $(Y,\gamma,0)$-DW-superprocess with $\gamma(\cdot)\equiv\gamma>0$ constant, (PC) holds, and $\big(H,(\mathbb{Q}_{\tau,m})_{(\tau,m)\in\hat{E}}\big)$ is the corresponding historical process on their canonical path spaces, Ω_X and Ω_H, respectively.

Theorem III.1.1. (Historical Cluster Representation). Let $m\in M_F^\tau(D)$ and $\tau\le s<t$. If $r_s(H_t)(\cdot)=H_t(\{y:y^s\in\cdot\})$, then

$$(III.1.1)\qquad \mathbb{Q}_{\tau,m}\left(r_s(H_t)\in\cdot\mid \mathcal{F}^H[\tau,s+]\right)(\omega)\overset{\mathcal{D}}{\equiv}\sum_{i=1}^{M}e_i\delta_{y_i}\,,$$

where $(e_i,y_i)_{i\le M}$ are the points of a Poisson point process on $\mathbb{R}_+\times D$ with intensity $\big(\frac{2}{\gamma(t-s)}\big)\big(\nu_{t-s}\times H_s(\omega)\big)$ and ν_{t-s} is an exponential law with mean $\gamma(t-s)/2$. That is, the right-hand side of (III.1.1) defines a regular conditional distribution for the random measure on the left side given $\mathcal{F}^H[\tau,s+]$.

Proof. By the Markov property of H we may assume $s=\tau$. Fix $m\in M_F^\tau(D)$. Let $A_1,\dots,A_n$ be a Borel partition of D and define $m_i=m(\cdot\cap A_i)$. Let $H^1,\dots,H^n$ be independent $(\mathcal{F}_t)$-historical processes with H^i starting at (τ,m_i) on some $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$. Then by checking the historical martingale problem $(HMP)_{\tau,m}$ we can easily see that $H\equiv\sum\limits_{i=1}^{n}H^i$ is an $(\mathcal{F}_t)$-historical process starting at (τ,m) and so has law $\mathbb{Q}_{\tau,m}$. For each i, the mean value property for historical processes (II.8.5) implies (recall $y=y^\tau$ m-a.e.) for each $t>\tau$

$$\mathbb{P}\left(H_t^i\left(\{y:y^\tau\in A_i^c\}\right)\right)=\int 1_{A_i}(y^\tau)P_{\tau,y^\tau}(y^\tau\in A_i^c)dm(y)=0.$$

The process inside the expected value on the left side is a.s. continuous in t by Exercise II.8.2 and so is identically 0 for all $t\ge\tau$ a.s. This implies

$$(III.1.2)\qquad H_t^i(\cdot)=H_t\left(\cdot\cap\{y^\tau\in A_i\}\right)\qquad i=1\dots n\quad\text{for all}\quad t\ge\tau\quad\text{a.s.}$$

It follows that if we start with $H.$ under $\mathbb{Q}_{\tau,m}$ and define H^i by (III.1.2), then

$$(III.1.3)\qquad \begin{array}{l}(H^i_\cdot,i\le n)\quad\text{are independent}\quad\mathcal{F}^H[\tau,t+]\text{-historical processes}\\ \text{starting at}\quad(\tau,m_i)_{i\le n},\quad\text{respectively.}\end{array}$$

In particular $\{H^i_{s+\tau}(1),s\ge 0\}_{i=1\dots n}$ are independent Feller diffusions (i.e. solutions of (II.5.10) with $g=0$) with initial values $\{m(A_i):i=1\dots n\}$, respectively. Recalling the Laplace transforms of $H_t^i(1)$ from (II.5.11), we have for any $\lambda_1,\dots,\lambda_n\ \ge 0$

and $f(y)=\sum\limits_{i=1}^{n}\lambda_i 1_{A_i}(y)$ $(y\in D^\tau)$,

$$\begin{aligned}\mathbb{Q}_{\tau,m}\Big(\exp\Big\{-\int f(y^\tau)H_t(dy)\Big\}\Big) &= \mathbb{Q}_{\tau,m}\Big(\exp\Big\{-\sum_{i=1}^{n}\lambda_i H_t^i(1)\Big\}\Big)\\ &= \exp\Big\{-\sum_{i=1}^{n}\frac{2\lambda_i m(A_i)}{2+\lambda_i(t-\tau)\gamma}\Big\}\\ &= \exp\Big\{-\int 2f(y)(2+f(y)(t-\tau)\gamma)^{-1}dm(y)\Big\}\\ &= E\Big(\exp\Big\{-\int ef(y)\Xi(de,dy)\Big\}\Big),\end{aligned}$$

where Ξ is a Poisson point process on $\mathbb{R}_+\times D$ with intensity $\frac{2}{\gamma(t-\tau)}\big(\nu_{t-\tau}\times m\big)$. As the above equation immediately follows for any Borel $f\ge 0$ on D, the proof is complete because we have shown $\int 1(y^\tau\in\cdot)H_t(dy)$ and $\int e1(y\in\cdot)\Xi(de,dy)$ have the same Laplace functionals. ∎

A consequence of the above argument is

Lemma III.1.2. If A is a Borel subset of $D(E)$, and $m\in M_F^\tau(D)$, then $X_s=H_{s+\tau}\big(\{y : y^\tau\in A\}\big)$ is a Feller branching diffusion (a solution of (II.5.10) with $g=0$), and for $t>\tau$,

$$\mathbb{Q}_{\tau,m}\left(H_s\left(\{y:y^\tau\in A\}\right)=0\quad\forall s\ge t\right)=\exp\left\{-\frac{m(A)2}{\gamma(t-\tau)}\right\}.$$

Proof. This is immediate from (III.1.3) and the extinction probabilities found in (II.5.12). ∎

The above Theorem shows for $s<t$, $H_t\big(\{y:y^s\in\cdot\}\big)$ is a purely atomic measure. The reader should be able to see that (conditionally) this measure is a Poisson superposition of exponential masses directly from Kolmogorov's and Yaglom's theorems (Theorem II.1.1). If $\tau=s=0$, the above may also be easily derived from the corresponding canonical measure representation (II.7.11) for H_t and projecting it down to y_0. Note that the exponential masses come from the last assertion of Theorem II.7.2(iii). An extension of the above cluster decomposition which describes the future evolution of the descendants of these finite number of clusters will be proved using similar ideas in Section III.6 (see Theorem III.6.1 and Corollary III.6.2).

Until otherwise indicated assume $\big((X_t)_{t\ge0},(\mathbb{P}_\mu)_{\mu\in M_F(\mathbb{R}^d)}\big)$ is a super-Brownian motion with branching rate $\gamma>0$ (we write X is SBM(γ)). This means X is a $(B,\gamma,0)$-DW-superprocess on its canonical space Ω_X, where B is a standard Brownian motion in $\mathbb{R}^d$ and γ is a positive constant. $H.$ will denote the corresponding historical process on Ω_H. We call H a historical Brownian motion (with branching rate γ).

The following result is essentially proved in Dawson-Iscoe-Perkins (1989) but first appeared in this form as Theorem 8.7 of Dawson-Perkins (1991). It gives a uniform modulus of continuity for all the paths in the closed support of H_t for all $t\ge 0$. The simple probabilistic proof given below seems to apply in a number of different settings. See, for example, Mueller and Perkins (1992) where it is applied to the supports of the solutions of a class of parabolic stochastic pde's. It also extends readily to more general branching mechanisms (see Dawson and Vinogradov (1994)) and to the interacting models considered in Chapter V (see Chapter 3 of Perkins (1995)).

Notation: $S(\mu)$ denotes the closed support of a measure μ on the Borel sets of a metric space. $h(r) = \big(r \log 1/r\big)^{1/2}$ is Lévy's modulus function.

Theorem III.1.3. (Historical Modulus of Continuity). If $\delta, c > 0$, let $K(\delta, c) = \{y \in C(\mathbb{R}^d) : |y(r) - y(s)| \le ch(|r-s|)\ \forall r, s \ge 0 \text{ satisfying } |r-s| \le \delta\}$. $\mathbb{Q}_{0,m}$ denotes the law of historical Brownian motion with branching rate $\gamma > 0$ starting at $(0, m)$.

(a) If $c > 2$, then $\mathbb{Q}_{0,m}$-a.s. there is a $\delta(c, \omega) > 0$ such that

$$S(H_t)(\omega) \subset K\big(\delta(c,\omega), c\big) \quad \forall t \ge 0.$$

Moreover there are constants $\rho(c) > 0$ and $C(d, c)$ such that

$$\mathbb{Q}_{0,m}(\delta(c) \le r) \le C(d,c) m(1) \gamma^{-1} r^{\rho} \text{ for } r \in [0,1].$$

(b) If $m \ne 0$ and $c < 2$, then $\mathbb{Q}_{0,m}$-a.s. for all $\delta > 0$ there is a t in $(0,1]$ such that $H_t\big(K(\delta,c)^c\big) > 0$.

Remark. This should be compared to Lévy's modulus of continuity for a simple Brownian path for which $c = \sqrt{2}$ is critical. This reflects the fact that the tree of Brownian paths underlying $H.$ has infinite length. We prove (a) below only for a sufficiently large c.

Proof of (a) (for large c). Use Lemma III.1.2 and the Markov property to see that

$$\begin{aligned}
&\mathbb{Q}_{0,m}\Big(H_t\Big(\Big\{y : \Big|y\Big(\frac{j}{2^n}\Big) - y\Big(\frac{j-1}{2^n}\Big)\Big| > ch(2^{-n})\Big\}\Big) > 0 \quad \exists t \ge (j+1)/2^n\Big) \\
&= \mathbb{Q}_{0,m}\Big(1 - \exp\Big\{ - \frac{2^{n+1}}{\gamma} H_{j/2^n}\Big(\Big\{y : \Big|y(\frac{j}{2^n}) - y\Big(\frac{j-1}{2^n}\Big)\Big| > ch(2^{-n})\Big\}\Big)\Big\}\Big) \\
&\le 2^{n+1}\gamma^{-1}\mathbb{Q}_{0,m}\left(H_{j/2^n}\left(|y(j/2^n) - y((j-1)/2^n)| > ch(2^{-n})\right)\right).
\end{aligned}$$

Now recall from (II.8.5) that the mean measure of H_t is just Wiener measure stopped at t. The above therefore equals

$$\begin{aligned}
&2^{n+1}\gamma^{-1} \int P_y\left(|B(j/2^n) - B((j-1)/2^n)| > ch(2^{-n})\right) dm(y) \\
&\le 2^{n+1}\gamma^{-1} m(1) c_d n^{d/2-1} 2^{-nc^2/2}
\end{aligned}$$

by a Gaussian tail estimate. Sum over $1 \le j \le n2^n$ to see that

$$\begin{aligned}
&\mathbb{Q}_{0,m}\Big(H_t\Big(\Big\{y : \Big|y\Big(\frac{j}{2^n}\Big) - y\Big(\frac{j-1}{2^n}\Big)\Big| > ch(2^{-n})\Big\}\Big) > 0, \\
&\qquad\qquad \text{for some } t \ge (j+1)2^{-n} \text{ and } 1 \le j \le n2^n\Big) \\
&\le c_d \gamma^{-1} m(1) n^{d/2} 2^{2n+1} 2^{-nc^2/2},
\end{aligned}$$

which is summable if $c > 2$. Assuming the latter, we may use Borel-Cantelli to see $\exists N(\omega) < \infty$ a.s. such that $H_t \equiv 0$ for $t \geq N(\omega)$ and

$$(III.1.4)\quad n \geq N(\omega) \Rightarrow \Big|y\Big(\frac{j}{2^n}\Big) - y\Big(\frac{j-1}{2^n}\Big)\Big| \leq ch(2^{-n}) \quad \forall j \geq 1,\ (j+1)2^{-n} \leq t, \qquad H_t\text{-a.a. y} \quad \forall t \leq n.$$

We now follow Lévy's proof for Brownian motion. Let $\delta(c_2, \omega) = 2^{-N(\omega)} > 0$ $a.s.$, where c_2 will be chosen large enough below. The required bound on $\mathbb{Q}_{0,m}(\delta(c_2) \leq r)$ is clear from the above bound and the extinction probability estimate formula (II.5.12). Let $N \geq t > 0$ and choose y outside of an H_t-null set so that (III.1.4) holds. Assume $r < s \leq t$ and $0 < s - r \leq 2^{-N}$ and choose $n \geq N$ so that $2^{-n-1} < s - r \leq 2^{-n}$. For $k \geq n$, choose $s_k \in \{j2^{-k} : j \in \mathbb{Z}_+\}$ such that $s_k + 2^{-k} \leq s$ and s_k is the largest such value (set $s_k = 0$ if $s < 2^{-k}$). One easily checks that

$$s_k \uparrow s,\ s_{k+1} = s_k + j_{k+1}2^{-(k+1)} \text{ for } j_{k+1} = 0, 1 \text{ or } 2 \text{ (and } j_{k+1} = 0 \text{ only arises if } s < 2^{-k-1}).$$

Note also that $s_k + (j_{k+1}+1)2^{-k-1} = s_{k+1} + 2^{-k-1} \leq s \leq t$. Therefore the choice of y and (III.1.4) imply $|y(s_{k+1}) - y(s_k)| \leq j_{k+1}ch(2^{-k-1})$, and so for some $c_1 > 0$ ($c_1 = 5c/\log 2$ will do),

$$(III.1.5)\quad |y(s) - y(s_n)| \leq \sum_{k=n}^{\infty} j_{k+1}ch(2^{-k-1}) \leq 2c\sum_{k=n}^{\infty} h(2^{-k-1}) \leq c_1h(2^{-n-1}) \leq c_1h(s-r).$$

Similarly one constructs $r_k \uparrow r$ so that

$$(III.1.6)\quad |y(r) - y(r_n)| \leq c_1h(s-r).$$

The restriction $s - r \leq 2^{-n}$ implies $s_n = r_n + j_n2^{-n}$ where $j_n = 0$ or 1 and so $s_n \leq s \leq t$, which means by (III.1.4) that

$$(III.1.7)\quad |y(s_n) - y(r_n)| \leq ch(2^{-n}) \leq c\sqrt{2}h(s-r).$$

(III.1.5)-(III.1.7) imply $|y(s) - y(r)| \leq (2c_1 + c\sqrt{2})h(s-r) \equiv c_2h(s-r)$. This proves $H_t\big(K(\delta(c_2), c_2)^c\big) = 0$ for $t \leq N(\omega)$ and so all $t \geq 0$ because $H_t = 0$ for $t > N$. $K(\delta, c_2)$ is closed and therefore $S(H_t) \subset K\big(\delta(c_2), c\big)$ for all $t \geq 0$.

To get (a) for any $c > 2$, one works with a finer set of grid points $\{(j + \frac{p}{M})\theta^n : j \in \mathbb{Z}_+,\ 0 \leq p < M\}$ ($\theta < 1$ sufficiently small and M large) in place of $\{j2^{-n} : j = 0, 1, 2\ldots\}$ to get a better approximation to r, s, as in Lévy's proof for Brownian motion. For example, see Theorem 8.4 of Dawson-Perkins (1991).

(b) Let $c < 2$ and $1 > \eta > 0$. If

$$B_j^n = \{y \in C : |y(2j2^{-n}) - y((2j-1)2^{-n})| > ch(2^{-n})\},$$

it suffices to show
(*III*.1.8)
$$\sup_{j\in\mathbb{N},(2j+1)2^{-n}<\eta} H_{(2j+1)2^{-n}}(B_j^n) > 0 \text{ for large } n \text{ for a.a. } \omega \text{ satisfying } \inf_{t<\eta} H_t(1) \geq \eta.$$

This is because the first event implies $\sup_{t<\eta} H_t(K(\delta,c)^c) > 0\ \forall\delta > 0$, and $\mathbb{Q}_{0,m}(\inf_{t<\eta} H_t(1) \geq \eta) \uparrow 1$ as $\eta \downarrow 0$. If

$$A_j^n = \{\omega : H_{(2j+1)2^{-n}}(B_j^n) = 0,\ H_{(2j+1)2^{-n}}(1) \geq \eta\},$$

then we claim that

$$\sum_{n=1}^{\infty} \mathbb{Q}_{0,m}(\cap_{1\leq j,(2j+1)2^{-n}\leq\eta} A_j^n) < \infty. \tag{III.1.9}$$

Assume this for the moment. The Borel-Cantelli Lemma implies

$$\begin{aligned}&\text{w.p.1 for } n \text{ large enough } \exists(2j+1)2^{-n} < \eta \text{ such that}\\ &H_{(2j+1)2^{-n}}(B_j^n) > 0 \text{ or } H_{(2j+1)2^{-n}}(1) < \eta,\end{aligned}$$

which implies

$$\text{w.p.1 either } \inf_{t<\eta} H_t(1) < \eta \text{ or for large } n \sup_{(2j+1)2^{-n}<\eta} H_{(2j+1)2^{-n}}(B_j^n) > 0.$$

This gives (III.1.8) and so completes the proof.

Turning to (III.1.9), note that

$$\begin{aligned}\mathbb{Q}_{0,m}(A_j^n|\mathcal{F}^H[0,2j2^{-n}+]) &\leq \mathbb{Q}_{0,m}(H_{(2j+1)2^{-n}}(B_j^n) = 0|\mathcal{F}^H[0,2j2^{-n}+])\\ &\leq \exp\Big\{-2^{n+1}\gamma^{-1}H_{2j2^{-n}}(B_j^n)\Big\},\end{aligned}$$

the last by Theorem III.1.1 with $t = (2j+1)2^{-n}$ and $s = 2j2^{-n}$. Let $R_{s,t}(y,\cdot)$ be the canonical measures associated with H, introduced in (II.8.6). Condition the above with respect to $\mathcal{F}^H[0,(2j-1)2^{-n}+]$ and use (II.8.6)(b) and the Markov property to see that

$$\begin{aligned}&\mathbb{Q}_{0,m}(A_j^n|\mathcal{F}^H[0,(2j-1)2^{-n}+])\\ &\quad\leq \exp\Big\{-\iint 1 - \exp(-2^{n+1}\gamma^{-1}\nu(B_j^n))R_{(2j-1)2^{-n},2j2^{-n}}(y,d\nu)H_{(2j-1)2^{-n}}(dy)\Big\}\\ &\quad\leq \exp\Big\{-\iint 1(\nu(1)\leq 2^{-n})[1-\exp(-2^{n+1}\gamma^{-1}\nu(B_j^n))]\\ &\qquad\qquad\qquad R_{(2j-1)2^{-n},2j2^{-n}}(y,d\nu)H_{(2j-1)2^{-n}}(dy)\Big\}\\ &\quad\leq \exp\Big\{-c_0\iint 1(\nu(1)\leq 2^{-n})2^{n+1}\gamma^{-1}\nu(B_j^n)\\ &\qquad\qquad\qquad R_{(2j-1)2^{-n},2j2^{-n}}(y,d\nu)H_{(2j-1)2^{-n}}(dy)\Big\},\end{aligned}$$

where $c_0 = c_0(\gamma) > 0$ satisfies $1 - e^{-2x/\gamma} \ge c_0 x$ for all $x \in [0,1]$. (II.8.6)(a) shows that on A^n_{j-1}

$$\begin{aligned}
&\mathbb{Q}_{0,m}(A^n_j|\mathcal{F}^H[0,(2j-1)2^{-n}+]) \\
&\quad\le \exp\Big\{-c_0 2^{n+1}\gamma^{-1}\int_0^\infty 1(\gamma 2^{-n-1}z \le 2^{-n})ze^{-z}dz \\
&\qquad\times \int P_{(2j-1)2^{-n},y}(|B(2j2^{-n}) - B((2j-1)2^{-n}| > ch(2^{-n}))H_{(2j-1)2^{-n}}(dy)\Big\} \\
&\quad\le \exp\Big\{-c_0 2^{n+1}\gamma^{-1}c_1 P^0(|B(1)| \ge c\sqrt{n\log 2})H_{(2j-1)2^{-n}}(1)\Big\} \\
&\quad\le \exp\{-c_2(\gamma,c)\eta 2^{n(1-c^2/2)}n^{-1/2}\},
\end{aligned}$$

for some $c_2 > 0$, at least for $n \ge n_0(c)$ by a Gaussian tail estimate and the fact that $H_{(2j-1)2^{-n}}(1) \le \eta$ on A^n_{j-1}. Therefore for $n \ge n_1(c,\eta)$,

$$\mathbb{Q}_{0,m}\Big(\cap_{j\ge 1,(2j+1)2^{-n}\le\eta} A^n_j\Big) \le \exp\Big\{\frac{-c_2\eta^2}{2}\frac{2^{n(2-c^2/2)}}{\sqrt{n}}\Big\},$$

which is summable over n since $c < 2$. This gives (III.1.9) and we are done. ∎

Remark. It is easy to obtain versions of the above result for continuous spatial motions other than Brownian motion (see Theorem 8.6 in Dawson-Perkins (1991)).

Notation. $\Pi_t : D(E) \to E$ is the projection map $\Pi_t(y) = y(t)$.

Recall X denotes SBM(γ) and H is the associated historical Brownian motion.

Corollary III.1.4. (a) $S(H_t)$ is compact in $C(\mathbb{R}^d)$ $\forall t > 0$ $\mathbb{Q}_{0,m}$-a.s.
(b) $S(X_t) = \Pi_t\big(S(H_t)\big)$ and hence is compact in $\mathbb{R}^d$ $\forall t > 0$ $\mathbb{P}_m$-a.s.

Proof. (a) Lemma III.1.2 shows that for any $\eta > 0$,

$$\begin{aligned}
\mathbb{Q}_{0,m}\left(H_s\big(\{|y_0| > R\}\big) = 0 \text{ for all } s \ge \eta\right) &= \exp\left\{-\frac{2m(\{(y_0) > R\})}{\gamma\eta}\right\} \\
&\to 1 \text{ as } R \to \infty.
\end{aligned}$$

This and the previous theorem show that for $\mathbb{Q}_{0,m}$-a.a. ω there is a $\delta(3,\omega) > 0$ and an $R(\omega) < \infty$ such that

$$S(H_t)(\omega) \subset K\big(\delta(3,\omega),3\big) \cap \{y : |y_0| \le R(\omega)\} \quad \forall t \ge \eta.$$

The set on the righthand side is compact by the Arzela-Ascoli theorem. Let $\eta \downarrow 0$ to complete the proof of (a).
(b) From (II.8.4) we may assume that $X_t = H_t \circ \Pi_t^{-1}$ for all $t \ge 0$. Note that $S(H_t) \subset \Pi_t^{-1}\Big(\Pi_t\big(S(H_t)\big)\Big)$ and therefore

$$X_t\left(\Pi_t\big(S(H_t)\big)^c\right) = H_t\left(\Pi_t^{-1}\left(\Pi_t\big(S(H_t)\big)\right)^c\right) \le H_t\big(S(H_t)^c\big) = 0.$$

This shows $\Pi_t\big(S(H_t)\big)$ supports X_t and as it is compact $\forall t > 0$ by (a), $S(X_t) \subset \Pi_t\big(S(H_t)\big)$ $\forall t > 0$ and is also compact a.s. If $w \in S(H_t)$, $H_t\big(\{y : |y_t - w_t| < \varepsilon\}\big) > 0$ for all $\varepsilon > 0$ and so $w_t \in S(X_t)$. This shows the reverse inclusion. ∎

A measure-valued process $X.$ has the compact support property (CSP) iff $S(X_0)$ compact implies $S(X_t)$ is compact for all $t > 0$ a.s. Corollary III.1.4 shows that super-Brownian motion has the (CSP) (in fact $S(X_0)$ need not be compact). Obviously this property fails for the heat kernel $P_t\phi$. The (CSP) for SBM was first proved by Iscoe (1988). The next result provides the natural rate of propagation for X suggested by the historical modulus of continuity.

Notation. $A \subset \mathbb{R}^d$, $\delta > 0$, $A^\delta = \{x \in \mathbb{R}^d : d(A,x) \equiv \inf\{|y-x| : y \in A\} < \delta\}$.

Corollary III.1.5. With probability 1 for any $c > 2$ there is a $\delta(c,\omega) > 0$ such that if $0 < t - s < \delta(c,\omega)$, then $S(X_t) \subset S(X_s)^{ch(t-s)}$.

To avoid an unexpected decline in $S(X_s)$ on the right side of this inclusion we need a lemma.

Lemma III.1.6. For $\mathbb{Q}_{0,m}$-a.a. ω if $0 \le s \le t$ and $y \in S(H_t)$, then $y(s) \in S(X_s)$.

Proof. If $0 \le s < s'$ are fixed, Theorem III.1.1 shows that conditional on $\mathcal{F}^H[0, s+]$, $H_{s'}\big(y(s) \in \cdot\big)$ is supported on a finite number of points $x_1 \ldots x_n$ in $S(X_s)$. The Markov property and (III.1.3) show that conditional on $\mathcal{F}^H[0, s'+]$,

$$\{H_t(y_s = x_i) : t \ge s'\}_{i \le n} \text{ and } \{H_t(y_s \notin \{x_1 \ldots x_n\}) : t \ge s'\}$$

are independent Feller diffusions. The latter is therefore a.s. identically 0 and so w.p.1 for all $t \ge s'$,

$$\begin{aligned}\{y(s) : y \in S(H_t)\} &\subset S\big(H_t(y_s \in \cdot)\big) \qquad \text{(trivial)}\\ &\subset S\big(H_{s'}(y_s \in \cdot)\big)\\ &\subset S(X_s).\end{aligned}$$

Take the union over all $0 \le s \le s'$ in $\mathbb{Q}$ to conclude

$$(III.1.8) \quad w.p.1 \quad \text{for all} \quad s \in \mathbb{Q}^{\ge 0} \quad \text{and all} \quad t > s \quad \{y(s) : y \in S(H_t)\} \subset S(X_s).$$

A simple consequence of our modulus of continuity and $X_t = H_t(y_t \in \cdot)$ is that if $B = B(x,\varepsilon)$, $B' = B(x,\varepsilon/2)$ and $m(B) = 0$, then $\mathbb{Q}_{\tau,m}$-a.s. $\exists \eta > 0$ such that $X_t(B') = 0$ for all $\tau \le t < \eta$. Use this and the strong Markov property at time $T_r(B) = \inf\{t \ge r : X_t(B) = 0\}$ where $B = B(x,\varepsilon)$ is a rational ball ($x \in \mathbb{Q}^d, \varepsilon \in \mathbb{Q}^{>0}$) and B' is as above to conclude:

$$(III.1.9) \qquad \begin{array}{l} w.p.1 \text{ for all } r \in \mathbb{Q}^{\ge 0} \text{ and rational ball } B\ \exists \eta > 0 \text{ such that} \\ X_s(B') = 0 \text{ for all } s \text{ in } [T_r(B), T_r(B) + \eta). \end{array}$$

Choose ω so that (III.1.8) and (III.1.9) hold. Let $y \in S(H_t)$, $s < t$ and suppose $y(s) \notin S(X_s)$ (the $s = t$ case is handled by Corollary III.1.4). Choose a rational ball B so that $y(s) \in B'$ and $X_s(B) = 0$, $\eta' > 0$, and a rational r in $(s - \eta', s]$. Then

$T_r(B) \le s$ because $X_s(B) = 0$ and so by (III.1.9) there is an open interval I in $(s-\eta', s+\eta')$ such that $X_u(B') = 0$ for all u in I. In particular there are rationals $u_n \to s$ such that $X_{u_n}(B') = 0$. On the other hand by (III.1.8) and the continuity of y, $y(u_n) \in B'$ and $y(u_n) \in S(X_{u_n})$ for n large which implies $X_{u_n}(B') > 0$, a contradiction. ∎

Proof of Corollary III.1.5. Apply Theorem III.1.3, Corollary III.1.4 and Lemma III.1.6 to see that w.p.1 if $0 < t - s < \delta(c,\omega)$,

$$S(X_t) = \Pi_t\big(S(H_t)\big) \subset \Pi_s\big(S(H_t)\big)^{ch(t-s)} \subset S(X_s)^{ch(t-s)}. \quad ∎$$

Remark. Presumably $c = 2$ is also sharp in Corollary III.1.5 if $d \ge 2$, although this appears to be open. It would be of particular interest to find the best result in $d = 1$ as the behaviour of $\partial S(X_t)$ in $d = 1$ could shed some light on the SPDE for super-Brownian motion in one dimension.

Definition. For $I \subset \mathbb{R}_+$ we call $\mathcal{R}(I) = \bigcup_{t\in I} S(X_t)$, the range of X on I, and $\overline{\mathcal{R}}(I) = \overline{\mathcal{R}(I)}$ is the closed range of X on I. The range of X is $\mathcal{R} = \bigcup_{\delta>0} \overline{\mathcal{R}}\big([\delta,\infty)\big)$.

It is not hard to see that $\mathcal{R} - \mathcal{R}(0,\infty)$ is at most a countable set of "local extinction points" (see Proposition 4.7 of Perkins (1990) and the discussion in Section III.7 below). $\mathcal{R}$ is sometimes easier to deal with than $\mathcal{R}((0,\infty))$. The reason for not considering $\mathcal{R}\big([0,\infty)\big)$ or $\overline{\mathcal{R}}\big([0,\infty)\big)$ is that it will be $\mathbb{R}^d$ whenever $S(X_0) = \mathbb{R}^d$.

Corollary III.1.7. $\overline{\mathcal{R}}\big([\delta,\infty)\big)$ is compact for all $\delta > 0$ a.s. $\overline{\mathcal{R}}\big([0,\infty)\big)$ is a.s. compact if $S(X_0)$ is.

Proof. Immediate from Corollaries III.1.4 and III.1.5. ∎

In view of the increase in the critical value of c in Theorem III.1.3 from that for a single Brownian path, it is not surprising that there are diffusions Y for which the (CSP) fails for the associated DW-superprocess X. Example 8.16 in Dawson-Perkins (1991) gives a time-inhomogeneous $\mathbb{R}$-valued diffusion and $T > 0$ for which $S(X_T) = \phi$ or $\mathbb{R}$ a.s. For jump processes instantaneous propagation is to be expected as is shown in the next Section.

2. Support Propagation and Super-Lévy Processes

Let Y be a Poisson (rate $\lambda > 0$) process on $\mathbb{Z}_+$ and consider X, the $(Y,\gamma,0)$-DW superprocess with $\gamma > 0$ constant. Then $A\phi(i) = \lambda\big(\phi(i+1) - \phi(i)\big)$, $i \in \mathbb{Z}_+$, and taking $\phi(i) = 1(i = j)$ in the martingale problem for X, we see that $X_t = \big(X_t(j)\big)_{j\in\mathbb{Z}_+}$ may also be characterized as the pathwise unique solution of

$$(III.2.1)\ \ X_t(j) = X_0(j) + \lambda\int_0^t \big(X_s(j-1) - X_s(j)\big)ds + \int_0^t \sqrt{\gamma X_s(j)}\,dB_s^j,\ j \in \mathbb{Z}_+.$$

Here $\{B^j : j \in \mathbb{Z}_+\}$ is a collection of independent linear Brownian motions, and $X_s(-1) \equiv 0$. Pathwise uniqueness holds by the method of Yamada and Watanabe (1971).

Let $X_0 = \alpha\delta_0$. Note that $X_t\big(\{0,\ldots,n\}\big)$ is a non-negative supermartingale and so sticks at 0 when it first hits 0 at time ζ_n. Evidently

$$\zeta_n \uparrow \zeta = \inf\{t : X_t(\mathbb{Z}_+) = 0\} < \infty \quad \text{a.s.}$$

Clearly $m_t = \inf S(X_t) = n$ if $t \in [\zeta_{n-1}, \zeta_n)$ ($\zeta_{-1} = 0$) and so is increasing in t a.s., and X_t becomes extinct as the lower end of its support approaches infinity. On the other hand it is clear (at least intuitively) from (III.2.1) that the mass at m_t will immediately propagate to $m_t + 1$ which in turn immediately propagates to $m_t + 2$, and so on. Therefore we have

$$(III.2.2) \qquad S(X_t) = \{m_t, m_t + 1, \ldots\} \quad \text{a.s.} \quad \forall t > 0.$$

This result will be a special case of Theorem III.2.4 below. Note, however, that $S(X_t) = \{m_t\}$ at exceptional times is a possibility by a simple comparison argument with the square of a low-dimensional Bessel process (see Exercise III.2.1).

The above example suggests $S(X_t)$ will propagate instantaneously to any points to which Y can jump. This holds quite generally (Corollary 5.3 of Evans-Perkins (1991)) but for the most part we consider only d-dimensional Lévy processes here. Our first result, however, holds for a general (Y, γ, g)-DW superprocess X with law $\mathbb{P}_{X_0}$. Recall if $Z_t = Z_0 + M_t + V_t$ is the canonical decomposition of a semimartingale then its local time at 0 is

$$(LT) \qquad L_t^0(Z) = \lim_{\varepsilon \downarrow 0} \varepsilon^{-1} \int_0^t 1(0 \le Z_s \le \varepsilon)\, d\langle M\rangle_s \quad \text{a.s.}$$

and in particular

$$(III.2.3) \qquad \int_0^\infty 1(Z_s = 0)\, d\langle M\rangle_s = 0.$$

Theorem III.2.1. If $\phi \in \mathcal{D}(A)_+$, then with probability 1 for Lebesgue-a.a. s, $X_s(A\phi) > 0$ implies $X_s(\phi) > 0$.

Proof. This should be intuitively clear from $(MP)_{X_0}$ as the drift $X_s(A\phi)ds > 0$ should keep $X_s(\phi) > 0$ a.s. Since $X_t(\phi) = X_0(\phi) + M_t(\phi) + \int_0^t X_s(A^g\phi)ds$, we have with probability one,

$$\begin{aligned}
L_t \equiv L_t^0\big(X(\phi)\big) &= \lim_{\varepsilon\downarrow 0}\int_0^t 1\big(0 < X_s(\phi) \le \varepsilon\big) X_s(\gamma\phi^2)ds\varepsilon^{-1} \\
&\le \|\gamma\phi\| \lim_{\varepsilon\downarrow 0}\int_0^t 1\big(0 < X_s(\phi) \le \varepsilon\big) X_s(\phi)\varepsilon^{-1}ds \\
&\le \|\gamma\phi\| \lim_{\varepsilon\downarrow 0}\int_0^t 1\big(0 < X_s(\phi) \le \varepsilon\big) ds \\
&= 0 \quad \text{a.s.}
\end{aligned}$$

Tanaka's Formula implies

$$\begin{aligned}
X_t(\phi)^+ = X_0(\phi) + \int_0^t 1\left(X_s(\phi) > 0\right) dM_s(\phi) + \int_0^t 1\left(X_s(\phi) > 0\right) X_s(A\phi)ds \\
+ \int_0^t 1\left(X_s(\phi) > 0\right) X_s(g\phi)ds.
\end{aligned}$$

Clearly $\int_0^t 1(X_s(\phi) = 0)\ X_s(g\phi)\, ds = 0$, and by (III.2.3), $\int_0^t 1\big(X_s(\phi) = 0\big)dM_s(\phi) = 0$. The above therefore implies

$$\begin{aligned}
X_t(\phi) = X_t(\phi)^+ &= X_0(\phi) + M_t(\phi) + \int_0^t X_s(A^g\phi)ds - \int_0^t 1(X_s(\phi) = 0)X_s(A\phi)ds \\
&= X_t(\phi) - \int_0^t 1(X_s(\phi) = 0)X_s(A\phi)ds.
\end{aligned}$$

We conclude that $\int_0^t 1(X_s(\phi) = 0)X_s(A\phi)ds = 0\ \forall t \ge 0$ a.s. and the result follows. ∎

Assume now that Y is a Lévy process in $\mathbb{R}^d$ with Lévy measure ν. Then $\mathcal{D}(A)$ contains C_K^∞, the C^∞-functions on $\mathbb{R}^d$ with compact support. Let B be an open ball in $\mathbb{R}^d$ and choose $\phi \in (C_K^\infty)_+$ such that $\{\phi > 0\} = B$. Then for $x \notin B$, $A\phi(x) = \int \phi(x+y)\nu(dy)$ (see, e.g., Theorem IV.4.1 of Gihman and Skorokhod (1975), or Example II.2.4(b) when Y is an asymmetric α-stable process). This means that $X_s(B) = 0$ implies $X_s(A\phi) = X_s * \nu(\phi)$, where $*$ denotes convolution of measures. Theorem III.2.1 therefore implies w.p.1 $X_s * \nu(B) > 0$ implies $X_s(B) > 0$, for Lebesgue a.a. s. Taking a union over balls with rational radii and centers we conclude

$$(III.2.4) \qquad S(X_s * \nu) \subset S(X_s) \quad \text{for Lebesgue a.a. } s > 0 \quad \text{a.s.}$$

The "Lebesgue a.a. s" is a nuisance as we would like to verify this inclusion for a fixed $s>0$ (it is false for all $s>0$ simultaneously as Exercise III.2.1 shows). The following result allows us to do this and also has several other applications.

Theorem III.2.2. Let X be the (Y,γ,g)-DW-superprocess where $\gamma(\cdot)\equiv\gamma>0$ is constant and $g\in C_b(E)$. Let $\mu_1,\mu_2\in M_F(E)$. The following are equivalent:
(i) $\mu_1 P_s \ll \mu_2 P_t \quad \forall\, 0<s\le t$
(ii) $\mathbb{P}_{\mu_1}(X_s\in\cdot)\ll\mathbb{P}_{\mu_2}(X_t\in\cdot) \quad \forall\, 0<s\le t$
(iii) $\mathbb{P}_{\mu_1}(X_{s+\cdot}\in\cdot)\ll\mathbb{P}_{\mu_2}(X_{t+\cdot}\in\cdot)$ (on $C(\mathbb{R}_+,M_F(E))$) $\quad\forall 0<s\le t$.

The original proof in Evans-Perkins (1991) used exact moment measure calculations and Theorem II.7.2. A simpler argument using only the latter is given at the end of the Section.

Example III.2.3. Let X be a super-α-stable process, i.e. $g=0$, $\gamma(\cdot)=\gamma>0$ constant and Y is the symmetric α-stable process in Example II.2.4(b) (and so is Brownian motion if $\alpha=2$). For any $\mu_1,\mu_2\in M_F(\mathbb{R}^d)-\{0\}$, (i) is trivial as μ_1P_s is equivalent to Lebesgue measure for all $s>0$. Therefore $\mathbb{P}_{\mu_1}(X_s\in\cdot)$ and $\mathbb{P}_{\mu_2}(X_t\in\cdot)$ are equivalent measures on $M_F(\mathbb{R}^d)$ and $\mathbb{P}_{\mu_1}(X\in\cdot)$ and $\mathbb{P}_{\mu_2}(X\in\cdot)$ are equivalent measures on $(\Omega_X,\sigma(X_r:r\ge\delta))$ for any $\delta>0$. For $0<\alpha<2$ the first equivalence allows us to consider a fixed s in (III.2.4) and conclude $S(X_s*\nu)\subset S(X_s)$ a.s. $\forall s>0$. Recall that $\nu(dx)=c|x|^{-d-\alpha}dx$ and conclude

$$(III.2.5)\qquad\qquad S(X_s)=\phi \quad\text{or}\quad \mathbb{R}^d \quad \text{a.s.}\quad \forall s>0.$$

A similar application of Theorem III.2.2 easily gives (III.2.2) for super-Poisson processes. More generally we have

Theorem III.2.4. Let Y be a Lévy process on $\mathbb{R}^d$ with Lévy measure ν, let $\gamma(\cdot)\equiv\gamma>0$ be constant and let X be the $(Y,\gamma,0)$-DW-superprocess starting at X_0 under $\mathbb{P}_{X_0}$. If ν^{*k} is the k-fold convolution of ν with itself then

$$\bigcup_{k=1}^{\infty} S(\nu^{*k}*X_t)\subset S(X_t)\ \mathbb{P}_{X_0}-\text{a.s.}\quad \forall t>0,\quad X_0\in M_F(\mathbb{R}^d).$$

Proof. Choose X_0 in $M_F(\mathbb{R}^d)$ so that $X_0(A)=0$ iff A is Lebesgue null. Then $P^{X_0}(Y_t\in A)=0$ iff A is Lebesgue null and so just as in the α-stable case above, Theorem III.2.2 and (III.2.4) imply that if $\Lambda=\{\mu\in M_F(E):S(\nu*\mu)\subset S(\mu)\}$ then $\mathbb{P}_{X_0}(X_t\in\Lambda)=1$ $\forall t>0$. The cluster decomposition (II.7.11) implies that for each $t>0$,

$$(III.2.6)\qquad\qquad R_t(x_0,\Lambda^c)=0 \quad\text{for Lebesgue a.a. } x_0.$$

Let $\tau_y: M_F(\mathbb{R}^d)\to M_F(\mathbb{R}^d)$ be the translation map $\tau_y(\mu)(A)=\int 1_A(x+y)d\mu(y)$. Then $R_t(x_0,\tau_y^{-1}(\cdot))=R_t(x_0+y,\cdot)$, e.g., by Theorem II.7.2(i) and the translation invariance of X_t^N. Clearly $\tau_y^{-1}(\Lambda)=\Lambda$ and so (III.2.6) implies $R_t(x_0,\Lambda^c)=0$ for any $x_0\in\mathbb{R}^d$. Another application of the cluster decomposition (II.7.11) (use $\bigcup_{i=1}^{n} S(\mu_i)=S(\sum_1^n\mu_i)$) shows $S(\nu*X_t)\subset S(X_t)$ $\mathbb{P}_{X_0}$-a.s. for any $X_0\in M_F(\mathbb{R}^d)$. Iterate this to complete the proof. ∎

Remark. Dawson's Girsanov theorem (Theorem IV.1.6 below) immediately gives the above for a general non-zero drift g in $C_b(E)$.

It is interesting to compare Theorem III.2.4 with the following result of Tribe (1992).

Theorem III.2.5. Assume Y is a Feller process on a locally compact metric space E, X is the Y-DW-superprocess starting at $X_0 \in M_F(E)$ under $\mathbb{P}_{X_0}$, and $\zeta = \inf\{t : X_t(1) = 0\}$. Then there is a random point F in E such that

$$\mathbb{P}_{X_0}\big(F \in A \mid X_s(1), s \geq 0\big)(\omega) = P^{X_0}(Y_{\zeta(\omega)} \in A)/X_0(1) \quad \text{a.s.} \quad \forall A \in \mathcal{E}$$

and

$$\lim_{t\uparrow\zeta} \frac{X_t(\cdot)}{X_t(1)} = \delta_F \quad \text{a.s. in} \quad M_F(E).$$

Proof. See Exercise III.2.2 below. ∎

This says the final extinction occurs at a single "death point" F, which, in view of the independence assumptions underlying X, is to be expected. On the other hand Example IV.2.3 shows that for an α-stable superprocess $S(X_t) = \mathbb{R}^d$ for Lebesgue a.a. $t < \zeta$ a.s. because of the ability of the mass concentrating near F to propagate instantaneously. The following result of Tribe (1992) shows the support will collapse down to F at exceptional times at least for $\alpha < 1/2$.

Theorem III.2.6. Assume X is a α-stable-DW-superprocess with $\alpha < 1/2$. Let F be as in Theorem III.2.5. For a.a. ω there are sequences $\varepsilon_n \downarrow 0$ and $t_n \uparrow \zeta$ such that $S(X_{t_n}) \subset B(F, \varepsilon_n)$.

We close this Section with the

Proof of Theorem III.2.2. The implications (ii) $\Rightarrow$ (i) and (ii) $\Rightarrow$ (iii) are immediate by considering $\mathbb{P}_{\mu_1}\big(X_s(\cdot)\big)$ and using the Markov property, respectively. As (iii) $\Rightarrow$ (ii) is trivial, only (i) $\Rightarrow$ (ii) requires a proof. Dawson's Girsanov Theorem (Theorem IV.1.6 below) reduces this implication to the case where $g \equiv 0$ which we now assume.

Assume (i) and choose $0 < s \leq t$. Write $R_u(m, \cdot)$ for $\int R_u(x_0, \cdot)dm(x_0)$ and set $R_u^*(m, \cdot) = R_u(m, \cdot)/R_u(m, M_F(E) - \{0\})$, where $R_u(x_0, \cdot)$ are the canonical measures of X from Theorem II.7.2 and $m \in M_F(E) - \{0\}$.

The first step is to reduce the problem to

$$(III.2.7) \qquad R_s(\mu_1, \cdot) \ll R_t(\mu_2, \cdot) \quad \text{on } M_F(E) - \{0\}.$$

By (II.7.11), $\mathbb{P}_{\mu_1}(X_s \in \cdot)$ and $\mathbb{P}_{\mu_2}(X_t \in \cdot)$ are the laws of $\sum_{i=1}^{N_1} \nu_i^1$ and $\sum_{i=1}^{N_2} \nu_i^2$, respectively, where N_1 and N_2 are Poisson with means $2\mu_1(1)/\gamma s$ and $2\mu_2(1)/\gamma t$, respectively, and conditional on N_1, N_2, $\{\nu_i^1 : i \leq N_1\}$ and $\{\nu_i^2 : i \leq N_2\}$ are i.i.d. with law $R_s^*(\mu_1, \cdot)$ and $R_t^*(\mu_2, \cdot)$, respectively. (III.2.7) implies the n-fold product of $R_s^*(\mu_1, \cdot)$ will be absolutely continuous to the n-fold product of $R_t^*(\mu_2, \cdot)$. Therefore we can sum over the values of N_1 and N_2 to obtain $\mathbb{P}_{\mu_1}(X_s \in \cdot) \ll \mathbb{P}_{\mu_2}(X_t \in \cdot)$ as required.

Let Ξ^{r,ν_0} denote a Poisson point process on $M_F(E) - \{0\}$ with intensity $R_r(\nu_0, \cdot)$. If $0 < \tau < t$, then Exercise II.7.2 (b) and (II.7.11) show that

$$(III.2.8) \qquad R_t(\mu_2, \cdot) = \int \mathbb{P}\Big(\int \nu\, \Xi^{t-\tau,\nu_0}(d\nu) \in \cdot\Big) R_\tau(\mu_2, d\nu_0) \text{ as measures on } M_F(E) - \{0\}.$$

This and the fact that $R_{t-\tau}(\nu_0,1)=\frac{2\nu_0(1)}{\gamma(t-\tau)}$, show that

$$\begin{aligned}R_t(\mu_2,\cdot)&\geq \mathbb{P}\Big(\int \nu\,\Xi^{t-\tau,\nu_0}(d\nu)\in\cdot,\ \Xi^{t-\tau,\nu_0}(M_F(E)-\{0\})=1\Big)\,R_\tau(\mu_2,d\nu_0)\\ (III.2.9)\qquad &=\int \exp\Big(-\frac{2\nu_0(1)}{\gamma(t-\tau)}\Big)R_{t-\tau}(\nu_0,\cdot)\,R_\tau(\mu_2,d\nu_0).\end{aligned}$$

Assume now that B is a Borel subset of $M_F(E)-\{0\}$ such that $R_t(\mu_2,B)=0$. Then (III.2.9) implies that

$$0=\iint R_{t-\tau}(x_0,B)d\nu_0(x_0)\,R_\tau(\mu_2,d\nu_0).$$

Recall from Theorem II.7.2 (c) that the mean measure associated with $R_\tau(\mu_2,\cdot)$ is μ_2P_τ. Therefore the above implies that

$$R_{t-\tau}(x_0,B)=0\quad \mu_2P_\tau-\text{a.a. }x_0.$$

Now apply (i) to see that for $0\leq h<\tau$,

$$R_{t-\tau}(x_0,B)=0\quad \mu_1P_{\tau-h}-\text{a.a. }x_0.$$

Now reverse the above steps to conclude

$$0=\int R_{t-\tau}(\nu_0,B)R_{\tau-h}(\mu_1,d\nu_0),$$

and for $s>\varepsilon>0$, set $\tau=t-s+\varepsilon$ and $h=t-s$ to get

$$\int R_{s-\varepsilon}(\nu_0,B)R_\varepsilon(\mu_1,d\nu_0)=0.$$

Use this result in (III.2.9) with our new parameter values to see that

$$\begin{aligned}0&=\int \mathbb{P}\Big(\int \nu\,\Xi^{s-\varepsilon,\nu_0}(d\nu)\in B,\ \Xi^{s-\varepsilon,\nu_0}(M_F(E)-\{0\})=1\Big)R_\varepsilon(\mu_1,d\nu_0)\\ (III.2.10)\quad &\geq R_s(\mu_1,B)-\int \mathbb{P}\Big(\Xi^{s-\varepsilon,\nu_0}(M_F(E)-\{0\})\geq 2\Big)R_\varepsilon(\mu_1,d\nu_0),\end{aligned}$$

where in the last line we used (III.2.8) with our new parameter values and also the fact that $0\notin B$ so that there is at least one point in the Poisson point process $\Xi^{s-\varepsilon,\nu_0}$. The elementary inequality $1-e^{-x}-xe^{-x}\leq x^2/2$ for $x\geq 0$ and Theorem II.7.2 (iii) show that if $\gamma'=\gamma(s-\varepsilon)/2$, then the last term in (III.2.10) is

$$\begin{aligned}&\int\Big(1-\exp(-\nu_0(1)/\gamma')-(\nu_0(1)/\gamma')\exp(-\nu_0(1)/\gamma')\Big)R_\varepsilon(\mu_1,d\nu_0)\\ &=\int_0^\infty\Big(1-e^{-x/\gamma'}-\frac{x}{\gamma'}e^{-x/\gamma'}\Big)\Big(\frac{2}{\gamma\varepsilon}\Big)^2e^{-2x/\gamma\varepsilon}\,dx\\ &\leq\int_0^\infty\frac{x^2}{2\gamma'^2}\Big(\frac{2}{\gamma\varepsilon}\Big)^2e^{2x/\gamma\varepsilon}\,dx\\ &=\gamma'^{-2}\gamma\varepsilon/2.\end{aligned}$$

Use this bound in (III.2.10) to conclude that $R_s(\mu_1, B) \le \gamma'^{-2}\gamma\varepsilon/2$ and for any ε as above and hence $R_s(\mu_1, B) = 0$, as required. ∎

Exercise III.2.1. Let X_t be the super-Poisson process which satisfies (III.2.1) with $X_0 = \alpha\delta_0$ ($\alpha > 0$). Prove that $P\big(S(X_t) = \{0\}\ \exists t > 0\big) > 0$ and conclude that (III.2.2) is not valid for all $t > 0$ a.s.

Hints. By a simple scaling argument we may take $\gamma = 4$.

(i) Show that

$$X_t(0) = \alpha - \int_0^t \lambda X_s(0)ds + \int_0^t 2\sqrt{X_s(0)}dB_s^0$$

$$X'_t \equiv X_t\big(\{0\}^c\big) = \int_0^t \lambda X_s(0)ds + \int_0^t 2\sqrt{X'_s}dB'_s,$$

where (B^0, B') are independent linear Brownian motions.

(ii) Let $T_1 = \inf\{t : X_t(0) \le \lambda^{-1}\}$, $T_2 = \inf\{t : t > T_1, X_t(0) \notin \big(\frac{1}{2}\lambda^{-1}, \frac{3}{2}\lambda^{-1}\big)\} - T_1$ and let

$$Y_t = X'_{T_1} + \frac{3}{2}t + \int_{T_1}^{t+T_1} 2\sqrt{Y_s}dB'_s.$$

Y is the square of a 3/2-dimensional Bessel process and hits 0 a.s. (see V.48 of Rogers and Williams (1987)), $Y_\cdot$ and $X_{\cdot+T_1}(0)$ are conditionally independent given $\mathcal{F}_{T_1}$. Argue that Y hits 0 before T_2 with positive probability and use this to infer the result.

Exercise III.2.2. (Tribe (1992)). Let X be a Y-DW-superprocess (hence $g \equiv 0, \gamma$ constant) on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where Y is a Feller process on a locally compact metric space E and $X_0 \in M_F(E) - \{0\}$. This means P_t is norm continuous on $C_\ell(E)$ the space of bounded continuous functions with a finite limit at ∞. Let $E_\infty = E \cup \{\infty\}$ be the one-point compactification of E.

(a) Let $C_t = \int_0^t \frac{1}{X_s(1)}\, ds$ for $t < \zeta$. It is well-known and easy to show that $C_{\zeta-} = \infty$ (see Shiga (1990), Theorem 2.1). Let $D_t = \inf\{s : C_s > t\}$ for $t \ge 0$, $\tilde{Z}_t(\cdot) = X_{D_t}(\cdot)$ and $Z_t(\cdot) = \tilde{Z}_t(\cdot)/\tilde{Z}_t(1)$. Show that if $\phi \in \mathcal{D}(A)$, then

$$Z_t(\phi) = X_0(\phi)/X_0(1) + \int_0^t \tilde{Z}_s(A\phi)ds + N_t(\phi),$$

where $N_t(\phi)$ is a continuous $(\mathcal{F}_{D_t})$-martingale such that

$$\langle N(\phi)\rangle_t = \int_0^t Z_s(\phi^2) - Z_s(\phi)^2 ds.$$

(b) Show that $\int_0^\infty \tilde{Z}_s(|A\phi|)ds < \infty$ a.s. and then use this to prove that $N_t(\phi)$ converges a.s. as $t \to \infty$ for all ϕ in $\mathcal{D}(A)$. Conclude that $Z_t(\cdot) \stackrel{a.s.}{\to} Z_\infty(\cdot)$ in $M_F(E_\infty)$ for some $Z_\infty \in M_F(E_\infty)$.
(c) Prove that $Z_\infty = \delta_F$ a.s. for some random point F in E_∞ and hence conclude that $X_t(\cdot)/X_t(1) \stackrel{a.s.}{\to} \delta_F(\cdot)$ as $t \uparrow \zeta$.
Hint. Prove $\lim_{t\to\infty} Z_t(\phi^2) - Z_t(\phi)^2 = 0$ for all ϕ in $\mathcal{D}(A)$, by using the a.s. convergence of $N_t(\phi)$.
(d) Use Exercise II.5.4 to see that

$$\mathbb{P}_{X_0}\big(F \in A \mid X_s(1), s \geq 0\big)(\omega) = P^{X_0}(Y_{\zeta(\omega)} \in A)/X_0(1) \text{ a.s. for all } A \in \mathcal{B}(E_\infty),$$

and in particular $F \in E$ a.s.

3. Hausdorff Measure Properties of the Supports.
Definition. If $h : \mathbb{R}_+ \to \mathbb{R}$ is strictly increasing and continuous near 0 and $h(0) = 0$ (write $h \in \mathcal{H}$), the Hausdorff h-measure of $A \subset \mathbb{R}^d$ is

$$h - m(A) = \liminf_{\delta\downarrow 0} \Big\{ \sum_{i=0}^\infty h\big(\mathrm{diam}(B_i)\big) : A \subset \bigcup_1^\infty B_i,\ B_i \text{ balls with } \mathrm{diam}(B_i) \leq \delta \Big\}$$

The Hausdorff dimension of A is $\dim(A) = \inf\{\alpha : x^\alpha - m(A) < \infty\} (\leq d)$.

The first result gives a global (in time) upper bound on $S(X_t)$ for $d \geq 3$ which allows one to quickly understand the 2-dimensional nature of $S(X_t)$. Until otherwise indicated $(X_t, t \geq 0)$ is a SBM(γ) starting at $X_0 \in M_F(\mathbb{R}^d)$ under $\mathbb{P}_{X_0}$, P_t is the Brownian semigroup and A is its (weak) generator.

Proposition III.3.1. Let $\psi(r) = r^2(\log^+ 1/r)^{-1}$ and $d \geq 3$. Then

$$\psi - m\big(S(X_t)\big) < \infty\ \forall t > 0 \text{ a.s.}$$

and, in particular, $\dim S(X_t) \leq 2\ \forall t > 0$ a.s.
Proof. By the Historical Cluster Representation (Theorem III.1.1)

$$S\big(H_{j/2^n}(\{y : y((j-1)2^{-n}) \in \cdot\})\big) = \{x_1^{j,n} \dots, x_{M(j,n)}^{j,n}\},$$

where conditional on $\mathcal{F}^H_{(j-1)2^{-n}}, M(j,n)$ has a Poisson distribution with mean $2^{n+1}\gamma^{-1}H_{(j-1)2^{-n}}(1)$ $(j, n \in \mathbb{N})$. The Historical Modulus of Continuity (Theorem III.1.3) implies that for a.a. ω if $2^{-n} < \delta(\omega, 3)$ and $t \in [j/2^n, (j+1)/2^n]$, then

$$\begin{aligned} S(X_t) &\subset S(X_{j2^{-n}})^{3h(2^{-n})} \qquad \text{(Corollary III.1.5)} \\ &= \big[\Pi_{j2^{-n}}\big(S(H_{j2^{-n}})\big)\big]^{3h(2^{-n})} \qquad \text{(Corollary III.1.4(b))} \\ &\subset \bigcup_{i=1}^{M(j,n)} B\left(x_i^{j,n}, 6h(2^{-n})\right). \end{aligned} \tag{III.3.1}$$

(The historical modulus is used in the last line.) A simple tail estimate for the Poisson distribution gives

$$
\begin{aligned}
&P\left(\bigcup_{j=1}^{n2^n}\{M(j,n)>2^{n+2}\gamma^{-1}\big(H_{(j-1)2^{-n}}(1)+1\big)\}\right)\\
&\quad\le\sum_{j=1}^{n2^n}E\left(\exp\left\{-2^{n+2}\gamma^{-1}(H_{(j-1)2^{-n}}(1)+1)\right\}\exp\left\{H_{(j-1)2^{-n}}(1)2^{n+1}\gamma^{-1}(e-1)\right\}\right)\\
&\quad\le n2^n\exp\left(-2^{n+2}\gamma^{-1}\right).
\end{aligned}
$$

By the Borel-Cantelli Lemma w.p.1 for large enough n (III.3.1) holds, and for all $j2^{-n}\le n$,

$$
\sum_{i=1}^{M(j,n)}\psi\big(12h(2^{-n})\big)\le\sup_t\big(H_t(1)+1\big)2^{n+2}\gamma^{-1}\psi\big(12h(2^{-n})\big)\le c\sup_t\big(H_t(1)+1\big).
$$

This implies $\psi-m\big(S(X_t)\big)<\infty\quad\forall t>0$ a.s. because $X_t=0$ for t large enough. ∎

Remark III.3.2. By taking unions over $j\le 2^nK$ in (III.3.1) we see from the above argument that for $K\in\mathbb{N}$ and $2^{-n}<\delta(\omega,3)$

$$
\overline{\bigcup_{t\in[2^{-n},K]}S(X_t)}\subset\bigcup_{j=1}^{K2^n}\bigcup_{i=1}^{M(j,n)}\overline{B\big(x_i^{j,n},6h(2^{-n})\big)}
$$

and for n large enough this is a union of at most $K_0 2^{2n+2}\gamma^{-1}\sup_t\big(H_t(1)+1\big)$ balls of radius $6h(2^{-n})$. As $X_t=0$ for $t>K(\omega)$, this shows that $f-m(\overline{\mathcal{R}}([\delta,\infty)))<\infty$ $\forall\delta>0$ where $f(r)=r^4(\log 1/r)^{-2}$, and so $\dim\mathcal{R}\le 4$. A refinement of this result which in particular shows that $\mathcal{R}$ is Lebesgue null in the critical 4-dimensional case is contained in Exercise III.5.1 below. The exact results are described below in Theorem III.3.9.

In order to obtain more precise Hausdorff measure functions one must construct efficient coverings by balls of variable radius (unlike those in Proposition III.3.1). Intuitively speaking, a covering is efficient if the balls contain a maximal amount of X_t-mass for a ball of its radius. This suggests that the lim sup behaviour of $X_t\big(B(x,r)\big)$ as $r\downarrow 0$ is critical. The following result of Rogers and Taylor (1961) (see Perkins (1988) for this slight refinement) plays a central role in the proof of the exact results described below (see Theorems III.3.8 and III.3.9).

Proposition III.3.3. There is a $c(d)>0$ such that for any $h\in\mathcal{H}$, $K>0$ and $\nu\in M_F(\mathbb{R}^d)$,
(a) $\nu(A)\le Kh-m(A)$ whenever A is a Borel subset of

$$
E_1(\nu,h,K)=\{x\in\mathbb{R}^d:\overline{\lim_{r\downarrow 0}}\,\nu\big(B(x,r)\big)/h(r)\le K\}.
$$

(b) $\nu(A) \ge c(d)Kh - m(A)$ whenever A is a Borel subset of

$$E_2(\nu, h, K) = \{x \in \mathbb{R}^d : \overline{\lim_{r\downarrow 0}}\, \nu\big(B(x,r)\big)/h(r) \ge K\}.$$

We can use Proposition III.3.3 (a) to get a lower bound on $S(X_t)$ which complements the upper bound in Proposition III.3.1.

Notation. If $d \ge 3$, let $\overline{h}_d(r) = r^2 \log^+ \frac{1}{r}$ and define $\overline{h}_2(r) = r^2(\log^+ \frac{1}{r})^2$.

Theorem III.3.4. If $d \ge 2$ there is a $c(d) > 0$ such that for all $X_0 \in M_F(\mathbb{R}^d)$, $\mathbb{P}_{X_0}$-a.s.

$\forall \delta > 0$ there is an $r_0(\delta,\omega) > 0$ so that $\sup\limits_{x,t\ge\delta} X_t\big(B(x,r)\big) \le \gamma c(d)\overline{h}_d(r)\ \forall r \in (0, r_0)$.

This result is very useful when handling singular integrals with respect to X_t (e.g., see the derivation of the Tanaka formula in Barlow-Evans-Perkins (1991)). Before proving this, here is the lower bound on $S(X_t)$ promised above.

Corollary III.3.5. If $d \ge 2$, $X_t(A) \le \gamma c(d)\overline{h}_d - m\big(A \cap S(X_t)\big)$ $\forall$ Borel set A and $t > 0$ $\mathbb{P}_{X_0}$-a.s. $\forall X_0 \in M_F(\mathbb{R}^d)$. In addition, if $\zeta = \inf\{t : X_t = 0\}$, then $\dim S(X_t) = 2$ for $0 < t < \zeta$ $\mathbb{P}_{X_0}$-a.s. $\forall X_0 \in M_F(\mathbb{R}^d)$.

Proof. By the previous result we may apply Proposition III.3.3 (a) to the sets $A \cap S(X_t)$ for all Borel A, $t > 0$ to get the first inequality. This with Proposition III.3.1 together imply that for $d \ge 2$, $\dim S(X_t) = 2$ $\forall 0 < t < \zeta$ a.s. ∎

Notation. If $f \ge 0$ is Borel measurable, let $G(f,t) = \int_0^t \sup_x P_s f(x) ds$.

Lemma III.3.6. If f is a non-negative Borel function such that $G(f,t)\gamma/2 < 1$, then

$$\mathbb{P}_{X_0}\big(\exp\left(X_t(f)\right)\big) \le \exp\Big\{X_0(P_t f)(1 - \frac{\gamma}{2}G(f,t))^{-1}\Big\} < \infty.$$

Proof. Let $k(s) = \big(1 - \frac{\gamma}{2}G(f, t-s)\big)^{-1}$ and $\phi(s,x) = P_{t-s}f(x)k(s)$ $(t > 0, s \in (0,t])$. If $\varepsilon > 0$, we claim $\phi\,|_{[0,t-\varepsilon]\times\mathbb{R}^d} \in \mathcal{D}(\vec{A})_{t-\varepsilon}$. To see this note that $P_\varepsilon f \in \mathcal{D}(A)$ so that $\frac{\partial}{\partial s}P_{t-s}f = \frac{\partial}{\partial s}P_{t-\varepsilon-s}(P_\varepsilon f) = -P_{t-\varepsilon-s}(AP_\varepsilon f)$ is continuous on $\mathbb{R}^d$ and is bounded on $[0, t-\varepsilon] \times \mathbb{R}^d$. The same is true of $\frac{\partial\phi}{\partial s}(s,x)$, and clearly

$$P_{t-s}f = P_{t-s-\varepsilon}(P_\varepsilon f) \in \mathcal{D}(A)$$

implies $\phi(s,\cdot) \in \mathcal{D}(A)$ and $A\phi_s = k(s)P_{t-s-\varepsilon}(AP_\varepsilon f)$ is bounded on $[0, t-\varepsilon] \times \mathbb{R}^d$. By Proposition II.5.7, for $s < t$

$$\begin{aligned} Z_s &\equiv X_s(P_{t-s}f)k_s \\ &= X_0(P_t f)k_0 + \int_0^s \int P_{t-r}f(x)k_r M(dr,dx) + \int_0^s X_r(P_{t-r}f)\dot{k}_r dr. \end{aligned}$$

By Itô's lemma there is a continuous local martingale N_s with $N_0 = 0$ so that for $s < t$,

$$e^{Z_s} = \exp\big(X_0(P_t f)k_0\big) + N_t + \int_0^s e^{Z_r} X_r \left[(P_{t-r}f)\dot{k}_r + \frac{\gamma}{2}(P_{t-r}f)^2 k_r^2\right] dr.$$

Our choice of k shows the quantity in square brackets is less than or equal to 0. This shows that e^{Z_s} is a non-negative local supermartingale, and therefore a supermartingale by Fatou's Lemma. Fatou's lemma also implies

$$E(e^{Z_t}) \le \liminf_{s\uparrow t} E(e^{Z_s}) \le e^{Z_0},$$

which gives the result. ∎

Remark III.3.7. The above proof shows the Lemma holds for $f \in \mathcal{D}(A)_+$ for any BSMP Y satisfying (II.2.2), or for $f \in b\mathcal{B}(\mathbb{R}^d)_+$ and any BSMP Y satisfying (II.2.2) and $P_t : b\mathcal{B}(\mathbb{R}^d)_+ \to \mathcal{D}(A)$ for $t > 0$. (We still assume $g = 0$, γ constant.) In particular Lemma III.3.6 is valid for the α-stable-DW-superprocess. Schied (1996) contains more on exponential inequalities as well as the idea underlying the above argument.

Proof of Theorem III.3.4. Recall $h(r) = (r \log 1/r)^{1/2}$. We first control $X_t\big(B(x, h(2^{-n}))\big)$ using balls in

$$\mathcal{B}_n = \left\{B\big(x_0, (\sqrt{d}+4)h(2^{-n})\big) : x_0 \in 2^{-n/2}\mathbb{Z}^d \cap [-n,n]^d\right\}.$$

Here $\varepsilon\mathbb{Z}^d = \{\varepsilon n : n \in \mathbb{Z}^d\}$. Assume $2^{-n} < \delta(\omega, 3)$, where $\delta(\omega, 3)$ is as in the Historical Modulus of Continuity. If $x \in [-n,n]^d$, choose $x_0 \in 2^{-n/2}\mathbb{Z}^d \cap [-n,n]^d$ such that $|x_0 - x| < \sqrt{d}\, 2^{-n/2} \le \sqrt{d}\, h(2^{-n})$ if $n \ge 3$. Assume $j \in \mathbb{Z}_+$ and $j2^{-n} \le t \le (j+1)2^{-n}$. For H_t-a.a. y if $y(t) \in B\big(x, h(2^{-n})\big)$, then

$$\begin{aligned}|y(j2^{-n}) - x_0| &\le |y(j2^{-n}) - y(t)| + |y(t) - x| + |x - x_0|\\ &\le (4+\sqrt{d})h(2^{-n}).\end{aligned}$$

If $B = B\big(x_0, (4+\sqrt{d})h(2^{-n})\big) \in \mathcal{B}_n$, this shows that

$$\begin{aligned}X_t\big(B(x, h(2^{-n}))\big) = H_t\big(y(t) \in B(x, h(2^{-n}))\big) &\le H_t\left(\{y(j2^{-n}) \in B\}\right)\\ &\equiv M_{j2^{-n},B}(t - j2^{-n}).\end{aligned}$$

Lemma III.1.2 and the Markov property show that $M_{j2^{-n},B}(t)$ is a martingale and so $M_{j2^{-n}}(t) = \sup_{B\in\mathcal{B}_n} M_{j2^{-n},B}(t)$ is a non-negative submartingale. The above bound implies

$$\sup_{j2^{-n}\le t\le (j+1)2^{-n}} \sup_{x\in[-n,n]^d} X_t\big(B(x, h(2^{-n}))\big) \le \sup_{t\le 2^{-n}} M_{j2^{-n}}(t)\ \forall\, j \in \mathbb{Z}_+ \quad \text{whenever } 2^{-n} \le \delta(\omega, 3). \tag{III.3.2}$$

We now use the exponential bound in Lemma III.3.6 to bound the right side of the above. Let $\beta = \beta(d) = 1 + 1(d = 2)$ and set $\varepsilon_n = c_1 2^{-n} n^{1+\beta(d)}$ where c_1 will be chosen (large enough) below. An easy calculation shows that

$$(III.3.3) \quad G\big(B(x,r),t\big) \equiv G\left(1_{B(x,r)},t\right) \le c_2(d) r^2 \left[1 + 1(d=2)\left(\log(\sqrt{t}/r)\right)^+\right]$$

from which it follows that

$$(III.3.4) \qquad \sup_{B \in \mathcal{B}_n, t \le n} G(B,t) \le c_3(d) 2^{-n} n^{\beta(d)}.$$

If $\delta > 0$ and $\lambda_n = c_3^{-1}\gamma^{-1}2^{n-1}n^{-\beta(d)}$, then the weak maximal inequality for submartingales implies

(III.3.5)

$$\begin{aligned}
&\mathbb{P}_{X_0}\left(\sup_{\delta 2^n \le j < n2^n} \sup_{t \le 2^{-n}} M_{j2^{-n}}(t) \ge \varepsilon_n\right) \\
&\quad \le n2^n \sup_{\delta 2^n \le j < n2^n} e^{-\lambda_n \varepsilon_n} \mathbb{P}_{X_0}\left(e^{\lambda_n M_{j2^{-n}}(2^{-n})}\right) \\
&\quad \le n2^n |\mathcal{B}_n| \sup_{\delta 2^n \le j < n2^n, B \in \mathcal{B}_n} e^{-\lambda_n \varepsilon_n} \mathbb{P}_{X_0}\Big(\exp\left(X_{j/2^n}(B) 2\lambda_n (2 - \lambda_n \gamma 2^{-n})^{-1}\right)\Big),
\end{aligned}$$

where we have used Lemma III.1.2, the Markov property and Lemma III.3.6 with $f \equiv \lambda_n$. The latter requires $\lambda_n 2^{-n}\gamma/2 < 1$ which is certainly true for $n \ge n_0$. In fact for $n \ge n_0$ (III.3.5) is no more than

$$(III.3.6) \qquad \begin{aligned}
&n2^n|\mathcal{B}_n| \sup_{\delta 2^n \le j < n2^n} \sup_{B \in \mathcal{B}_n} e^{-\lambda_n \varepsilon_n} \mathbb{P}_{X_0}\left(\exp\left(2\lambda_n X_{j/2^n}(B)\right)\right) \\
&\quad \le n2^n|\mathcal{B}_n| e^{-\lambda_n \varepsilon_n} \sup_{\delta 2^n \le j < n2^n} \sup_{B \in \mathcal{B}_n} \exp\left(4\lambda_n P^{X_0}(B_{j/2^n} \in B)\right)
\end{aligned}$$

by another application of Lemma III.3.6, this time with $f = 2\lambda_n 1_B$. We also use (III.3.4) here to see that

$$2\lambda_n G(B, j/2^{-n}) \le 2\lambda_n c_3 2^{-n} n^{\beta(d)} = \gamma^{-1},$$

so that Lemma III.3.6 may be used and the upper bound given there simplifies to the expression in (III.3.6). An elementary bound shows that the right side of (III.3.6) is at most

$$\begin{aligned}
&c n^{1+d} 2^{n+nd/2} e^{-\lambda_n \varepsilon_n} \exp\left(4\lambda_n c(\delta, X_0(1)) h(2^{-n})^d\right) \\
&\quad \le c n^{1+d} 2^{n(1+d/2)} \exp\left(-c_1 c_3^{-1}\gamma^{-1}2^{-1}n\right) c'\big(\delta, X_0(1)\big),
\end{aligned}$$

which is summable if $c_1 = \gamma 2 c_3\big(1 + \frac{d}{2}\big) \equiv \gamma c'(d)$. Borel-Cantelli and (III.3.2) imply that a.s. for large enough n,

$$\sup_{\delta \le t \le n} \sup_{x \in [-n,n]^d} X_t\left(B(x, h(2^{-n}))\right) < \gamma c' 2^{-n} n^{1+\beta(d)} \le \gamma c'' \overline{h}_d\big(h(2^{-n})\big).$$

An elementary interpolation completes the proof. ∎

In view of Proposition III.3.1 and Corollary III.3.5 it is natural to ask if there is an exact Hausdorff measure function associated with $S(X_t)$.

Notation.

$$h_d(r) = \begin{cases} r^2 \log^+ \log^+ 1/r & d \ge 3 \\ r^2(\log^+ 1/r)(\log^+ \log^+ \log^+ 1/r) & d = 2. \end{cases} \qquad \bigl(\log^+ x = (\log x) \vee e^e\bigr)$$

Theorem III.3.8. Assume X is a SBM(γ) starting at μ under $\mathbb{P}_\mu$.
(a) $[d \ge 2, t$ fixed$]$ There is a universal constant $c(d) > 0$ such that $\forall \mu \in M_F(\mathbb{R}^d)$, $t > 0$

$$X_t(A) = \gamma c(d) h_d - m\bigl(A \cap S(X_t)\bigr) \quad \forall A \in \mathcal{B}(\mathbb{R}^d) \quad \mathbb{P}_\mu\text{-a.s.}$$

(b) $[d \ge 2, t$ variable$]$ There are universal constants $0 < c(d) \le C(d) < \infty$ such that for any $\mu \in M_F(\mathbb{R}^d)$

(i) If $d \ge 3$, $\gamma c(d) h_d - m\bigl(A \cap S(X_t)\bigr) \le X_t(A) \le \gamma C(d) h_d - m\bigl(A \cap S(X_t)\bigr)$ $\forall A \in \mathcal{B}(\mathbb{R}^d)$ $\forall t > 0$ $\mathbb{P}_\mu$-a.s.

(ii) If $d = 2$, $\gamma c(2) h_3 - m\bigl(A \cap S(X_t)\bigr) \le X_t(A) \le \gamma C(2) \overline{h}_2 - m\bigl(A \cap S(X_t)\bigr)$ $\forall A \in \mathcal{B}(\mathbb{R}^d)$ $\forall t > 0$ $\mathbb{P}_\mu$-a.s.

(c) $[d = 1]$ There is a jointly continuous process $\{u(t,x) : t > 0, x \in \mathbb{R}\}$ such that

$$X_t(dx) = u(t,x)dx \quad \forall t > 0 \quad \mathbb{P}_\mu\text{-a.s.}$$

Remarks. (1) (b) shows that if $d \ge 2$ then w.p.1 for any $t > 0$ $S(X_t)$ is a singular set of Hausdorff dimension 2 whenever it is non-empty. This fact has already been proved for $d \ge 3$ (Corollary III.3.5).

(2) (a) and (b) state that X_t distributes its mass over $S(X_t)$ in a deterministic manner. This extreme regularity of the local structure of $S(X_t)$ is due to the fact that for $d \ge 2$ the local density of mass at x is due entirely to close cousins of "the particle at x" and so will exhibit strong independence in x. The strong recurrence for $d = 1$ means this will fail in $\mathbb{R}^1$ and a non-trivial density, u, exists.

(3) We conjecture that one may take $c(d) = C(d)$ in (b)(i) for $d \ge 3$. The situation in the plane is much less clear.

(4) Curiously enough the exact Hausdorff measure functions for $S(X_t)$ are exactly the same as those for the range of a Brownian path (see Ciesielski-Taylor (1962), Taylor (1964)) although these random sets certainly look quite different. The two sets behave differently with respect to packing measure: $h(s) = s^2(\log\log 1/s)^{-1}$ is an exact packing measure function for the range of a Brownian path for $d \ge 3$ while $h(s) = s^2(\log 1/s)^{-1/2}$ is critical for $S(X_t)$ for $d \ge 3$ (see Le Gall-Perkins-Taylor (1995)).

(5) (b) is proved in Perkins (1988, 1989). In Dawson-Perkins (1991), (a) was then proved for $d \ge 3$ by means of a 0-1 law which showed that the Radon-Nikodym derivative of X_t with respect to $h_d - m\bigl(\,\cdot \cap S(X_t)\bigr)$ is a.s. constant. The more delicate 2-dimensional case in (a) was established in Le Gall-Perkins (1995) using the Brownian snake. This approach has proved to be a very powerful tool in the study of path properties of (and other problems associated with) DW-superprocesses. The $d = 1$ result was proved independently by Reimers (1989) and Konno-Shiga (1988) and will be analyzed in detail in Section III.4 below. The existence of a density at a fixed time was proved by Roelly-Coppoletta (1986). The first Hausdorff measure result, $\dim S(X_t) \le 2$, was established by Dawson and Hochberg (1979).

Here are the exact results for the range of super-Brownian motion, promised after Proposition III.3.1.

Theorem III.3.9. (a) $d \ge 4$. Let

$$\psi_d(r) = \begin{cases} r^4 \log^+ \log^+ 1/r & d > 4 \\ r^4(\log^+ 1/r)(\log^+ \log^+ \log^+ 1/r) & d = 4 \end{cases}.$$

There is a $c(d) > 0$ such that for all $X_0 \in M_F(\mathbb{R}^d)$,

$$\int_0^t X_s(A)ds = \gamma c(d)\psi_d - m\left(A \cap \mathcal{R}((0,t])\right) \;\forall A \in \mathcal{B}(\mathbb{R}^d) \;\; \forall t \ge 0 \; \mathbb{P}_{X_0} - \text{a.s.}$$

(b) $d \le 3$. Assume X_0 has a bounded density if $d = 2, 3$. Then there is a jointly continuous density $\{v(t,x) : t \ge 0, x \in \mathbb{R}^d\}$ such that

$$\int_0^t X_s(A)ds = \int_A v(t,x)dx \quad \forall A \in \mathcal{B}(\mathbb{R}^d) \quad \forall t \ge 0 \quad \mathbb{P}_{X_0} - \text{a.s.}$$

Discussion. (a) for $d > 4$ is essentially proved in Dawson-Iscoe-Perkins (1989). There upper and lower bounds with differing constants were given. Le Gall (1999) showed that by a $0-1$ law the above constants were equal and used his snake to derive the critical 4-dimensional result.

(b) is proved in Sugitani (1989). ∎

Consider now the analogue of Theorem III.3.8 when X is an α-stable-DW-superprocess ($g \equiv 0$, $\gamma(\cdot) \equiv \gamma > 0$). That is Y is the symmetric α-stable process in $\mathbb{R}^d$ considered in Example II.2.4 (b). Let

$$h_{d,\alpha}(r) = r^\alpha \log^+ \log^+ 1/r \quad \text{if} \quad d > \alpha$$
$$\overline{h}_{d,\alpha}(r) = r^\alpha (\log^+ 1/r)^2 \quad \text{if} \quad d = \alpha$$
$$\psi_{d,\alpha}(r) = \begin{cases} r^\alpha & \text{if } d > \alpha \\ r^\alpha \log^+ 1/r & \text{if } d = \alpha \end{cases}.$$

In Perkins (1988) it is shown that for $d > \alpha$ there are real constants $0 < c_1 \le c_2$, depending only on (d, α) so that

$$(III.3.7) \qquad \gamma c_1 \le \overline{\lim_{r \downarrow 0}} \frac{X_t\big(B(x,r)\big)}{h_{d,\alpha}(r)} \le \gamma c_2 \quad X_t\text{-a.a. } x \quad \forall t > 0 \quad \mathbb{P}_{X_0}\text{-a.s.}$$

Let $\Lambda_t(\omega)$ be the set of x for which the above inequalities hold. Then $\Lambda_t(\omega)$ is a Borel set supporting $X_t(\omega)$ for all $t > 0$ a.s. and Proposition III.3.3 shows that (for $d > \alpha$)

$$(III.3.8) \qquad \begin{aligned} \gamma c(d) c_1 h_{d,\alpha} - m(A \cap \Lambda_t) \le X_t(A) \le & \gamma c_2 h_{d,\alpha} - m(A \cap \Lambda_t) \\ & \forall A \in \mathcal{B}(R^d), \; t > 0 \;\; \mathbb{P}_{X_0}\text{-a.s.} \end{aligned}$$

For t fixed a $0-1$ law then shows the $\overline{\lim}$ in (III.3.7) is γc_3 for X_t-a.a. x, say for $x \in \Lambda_t'$, a.s. and $X_t(\cdot) = c_4 \gamma h_{d,\alpha} - m(\cdot \cap \Lambda_t')$ a.s. for some $c_3, c_4 > 0$ (see Theorem

5.5 of Dawson-Perkins (1991)). Analogous results are also shown for $d=\alpha$ (=1 or 2) with $\psi_{d,\alpha}-m$ in the lower bound and $\overline{h}_{d,\alpha}-m$ in the upper bound. Such results are clearly false for $\alpha<2$ if Λ_t is replaced by $S(X_t)(=\phi$ or $\mathbb{R}^d$ a.s. by Example III.2.3).

(III.3.7) suggests we define the *Campbell measure*, $Q_t \in M_F(M_F(\mathbb{R}^d)\times\mathbb{R}^d)$, associated with X_t by

$$Q_t(A\times B)=\int 1_A(X_t)X_t(B)d\mathbb{P}_{X_0}.$$

If $(X,Z)\in M_F(\mathbb{R}^d)\times\mathbb{R}^d$ are coordinate variables on $M_F(\mathbb{R}^d)\times\mathbb{R}^d$,then under Q_t, Z is chosen at random according to X. The regular conditional probabilities $Q_t(X\in\cdot\mid Z=x)$ are the Palm measures of X_t and describe X_t from the perspective of a typical point x in the support of X_t (see Dawson-Perkins (1991), Chapter 4, and Dawson (1992), Chapter 6, for more on Palm measures of superprocesses).

The first step in deriving (III.3.7) is to find the asymptotic mean size of $X_t\big(B(x,r)\big)$ when x is chosen according to X_t. (a) of the following Exercise is highly recommended.

Exercise III.3.1. Let X be the α-stable-DW-superprocess and $Q_t(dX,dZ)$ be the Campbell measure defined above. If Y is the α-symmetric stable process in $\mathbb{R}^d$, Y_t has a smooth symmetric density $p_t(y)=p_t(|x|)$ such that $p_1(\cdot)$ is decreasing on $[0,\infty)$ and $p_1(r)\le c(1+r)^{-(d+\alpha)}$.

(a) If $d\ge\alpha$, show there is a constant $k_{d,\alpha}>0$ such that

$$(III.3.9)\qquad \begin{aligned}\lim_{r\downarrow 0}\frac{E\left(\int X_t\big(B(x,r)\big)dX_t(x)\right)}{\psi_{d,\alpha}(r)}&\equiv\lim_{r\downarrow 0}\frac{\int X\big(B(Z,r)\big)Q_t(dX,dZ)}{\psi_{d,\alpha}(r)}\\&=k_{d,\alpha}\gamma X_0(1)\quad\forall t>0.\end{aligned}$$

Also show there is a $c=c\big(d,\alpha,\gamma,X_0(1)\big)$ such that for any $\delta>0$ there is an $r_0(\delta)>0$ satisfying

$$(III.3.10)\qquad E\left(\int X_t\big(B(x,r)\big)dX_t(x)\right)\le c\psi_{d,\alpha}(r)\quad\forall t\in[\delta,\delta^{-1}],\quad r\le r_0(\delta).$$

(b) Show the above results remain valid for $d<\alpha$ if $\psi_{d,\alpha}(r)=r$, $k_{d,\alpha}$ may depend on (t,γ,X_0), and c may also depend on δ.

(c) [Palm measure version] If $X_0P_t(x)=\int p_t(y-x)X_0(dy)$, show that

$$\begin{aligned}&Q_t\left(X\left(B(Z,r)\right)\mid Z=x_0\right)\\&\quad=P^{X_0}\left(Y_t\in B(x_0,r)\right)+\frac{\gamma E^{X_0}\Big(\int_0^t P^{Y_s}\left(|Y_{t-s}-x_0|<r\right)p_{t-s}(x_0-Y_s)ds\Big)}{X_0P_t(x_0)}.\end{aligned}$$

*(d) Show there is a constant $k_{d,\alpha}$ (which may also depend on (γ,t,x_0,X_0) if $d<\alpha$) such that for any $t>0$, $x_0\in\mathbb{R}^d$,

$$(III.3.11)\qquad \lim_{r\downarrow 0}\frac{Q_t\left(X\big(B(Z,r)\big)\mid Z=x_0\right)}{\psi_{d,\alpha}(r)}=\gamma k_{d,\alpha}.$$

Recall that $\psi_{d,\alpha}(r) = r$ if $d < \alpha$.

The above Exercise shows that if $d > \alpha$, the mean of $X_t\big(B(x,r)\big)$ when x is chosen according to X_t behaves like $k_{d,\alpha}\psi_{d,\alpha}(r)$ as $r \downarrow 0$. This explains the r^α part of $h_{d,\alpha}$ in (III.3.7). The $\log\log 1/r$ then comes from the exponential tail of $X_t\big(B(x,r)\big)$ and the fact that it suffices to consider $r \downarrow 0$ through a geometric sequence (each exponentiation produces a log). It is an easy Borel-Cantelli Exercise to use the above mean results to obtain weaker versions of the upper bounds in (III.3.7) and (III.3.8).

Exercise III.3.2. (a) Use (III.3.10) to show that if $d \geq \alpha$, then for any $\varepsilon > 0$, $t > 0$

$$\lim_{r\downarrow 0} \frac{X_t\big(B(x,r)\big)}{\psi_{d,\alpha}(r)(\log^+ 1/r)^{1+\varepsilon}} = 0 \quad X_t\text{-a.a. } x \;\; \mathbb{P}_{X_0}\text{-a.s.}$$

(b) If $\phi^\varepsilon_{d,\alpha}(r) = \psi_{d,\alpha}(r)(\log^+ 1/r)^{1+\varepsilon}$, show that $\forall t > 0$,

$$\phi^\varepsilon_{d,\alpha} - m(A) = 0 \quad \text{implies} \quad X_t(A) = 0 \quad \forall A \in \mathcal{B}(\mathbb{R}^d),\ \varepsilon > 0 \quad \mathbb{P}_{X_0}\text{-a.s.}$$

4. One-dimensional Super-Brownian motion and Stochastic PDE's

In this section we study super-Brownian motion in one dimension with constant branching rate $\gamma > 0$. In particular, we will prove Theorem III.3.8(c) and establish a one-to-one correspondence between the density of super-Brownian motion in one dimension and the solution of a parabolic stochastic pde driven by a white noise on $\mathbb{R}_+ \times \mathbb{R}$.

Let $\mathcal{F}_t$ be a right continuous filtration on $(\Omega, \mathcal{F}, P)$ and let $\mathcal{P} = \mathcal{P}(\mathcal{F}_\cdot)$ denote the σ-field of $\mathcal{F}_t$-predictable sets in $\mathbb{R}_+ \times \Omega$. Let $|A|$ denote the Lebesgue measure of a Borel set in $\mathbb{R}^d$ and let $\mathcal{B}_F(\mathbb{R}^d) = \{A \in \mathcal{B}(\mathbb{R}^d) : |A| < \infty\}$.

Definition. An $(\mathcal{F}_t)$-white noise, W, on $\mathbb{R}_+ \times \mathbb{R}^d$ is a random process $\{W_t(A) : t > 0, A \in \mathcal{B}_F(\mathbb{R}^d)\}$ such that

(*i*) $W_t(A \cup B) = W_t(A) + W_t(B)$ a.s. for all disjoint $A, B \in \mathcal{B}_F(\mathbb{R}^d)$ and $t > 0$.

(*ii*) $\forall A \in \mathcal{B}_F(\mathbb{R}^d),\ t \mapsto W_t(A)$ is an $(\mathcal{F}_t)$-Brownian motion starting at 0 with diffusion parameter $|A|$.

See Chapters 1 and 2 of Walsh (1986) for more information on white noises. Note that if W is above and A and B are disjoint sets in $\mathcal{B}_F$, then

$$2W_t(A)W_t(B) = W_t(A \cup B)^2 - W_t(A)^2 - W_t(B)^2$$

is an $(\mathcal{F}_t)$-martingale. It follows that W is an orthogonal martingale measure in the sense of Chapter 2 of Walsh (1986) and Proposition 2.10 of Walsh (1986) shows that the above definition is equivalent to that given in Walsh (1986). As in Section II.5 (see Chapter 2 of Walsh (1986)) one can define a stochastic integral $W_t(\psi) = \int_0^t \int \psi(s, \omega, x)dW(s, x)$ for

$$\psi \in \mathcal{L}^2_{\text{loc}}(W) \equiv \{\psi : \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \to \mathbb{R} : \psi \text{ is } \mathcal{P} \times \mathcal{B}(\mathbb{R}^d) - \text{measurable},$$
$$\int_0^t \int \psi(s,\omega,x)^2 dx\, ds < \infty\ \forall t > 0 \text{ a.s. } \}.$$

The map $\psi \mapsto W(\psi)$ is a linear map from $\mathcal{L}^2_{\text{loc}}$ to the space of continuous $\mathcal{F}_t$-local martingales and $W_t(\psi)$ has predictable square function

$$\langle W(\psi)\rangle_t = \int_0^t \int \psi(s,\omega,x)^2 dx\, ds \quad \text{for all } t > 0 \text{ a.s.}$$

Notation. If f, g are measurable functions on $\mathbb{R}^d$, let $\langle f, g\rangle = \int f(x)g(x)\,dx$ if the integral exists.

Definition. Let W be an $(\mathcal{F}_t)$-white noise on $\mathbb{R}_+ \times \mathbb{R}^d$ defined on $\bar\Omega = (\Omega, \mathcal{F}, \mathcal{F}_t, P)$, let $m \in M_F(\mathbb{R}^d)$ and let $f : \mathbb{R}_+ \to \mathbb{R}$. We say that an adapted continuous process $u : (0,\infty) \times \Omega \to C_K(\mathbb{R}^d)_+$ is a solution of

$$(SPDE)^f_m \qquad \frac{\partial u}{\partial t} = \frac{\Delta u}{2} + \sqrt{\gamma u}\dot W + f(u), \qquad u_{0+}(x)dx = m$$

on $\bar\Omega$ iff for every $\phi \in C^2_b(\mathbb{R}^d)$,

$$\langle u_t, \phi\rangle = m(\phi) + \int_0^t \langle u_s, \frac{\Delta\phi}{2}\rangle\, ds + \int_0^t \phi(x)\sqrt{\gamma u(s,x)}dW(s,x)$$
$$+ \int_0^t \int f(u(s,x))\phi(x)dx\, ds \quad \text{for all } t > 0 \text{ a.s.}$$

Remark III.4.1. Use the fact that $C^2_b(\mathbb{R}^d)$ contains a countable convergence determining class (such as $\{\sin(u\cdot x), \cos(u\cdot x) : u \in \mathbb{Q}^d\}$) to see that $(\text{SPDE})^f_m$ implies that $\lim_{t\to 0+} u_t(x)dx = m$ a.s. in $M_F(\mathbb{R}^d)$. We have been able to choose a rather restrictive state space for u_t because the Compact Support Property for SBM (Corollary III.1.4) will produce solutions with compact support. This property will persist for stochastic pde's in which the square root in the white noise integrand is replaced by any positive power less than 1, but fails if this power equals 1 (see Mueller and Perkins (1992) and Krylov (1997)).

We write $(\text{SPDE})_m$ for $(\text{SPDE})^0_m$. The one-dimensional Brownian density and semigroup are denoted by p_t and P_t, respectively. We set $p_t(x) = 0$ if $t \le 0$.

Theorem III.4.2. (a) Let X be an $(\mathcal{F}'_t)$-SBM(γ) in one spatial dimension, starting at $X_0 \in M_F(\mathbb{R})$ and defined on $\bar\Omega' = (\Omega', \mathcal{F}', \mathcal{F}'_t, P')$. There is an adapted continuous $C_K(\mathbb{R})$-valued process $\{u_t : t > 0\}$ such that $X_t(dx) = u_t(x)\,dx$ for all $t > 0$ P'-a.s. Moreover for all $t > 0$ and $x \in \mathbb{R}$,

$$(III.4.1) \quad u_t(x) = P_tX_0(x) + \int_0^t \int p_{t-s}(y-x)dM(s,y) \quad P' - \text{a.s., and}$$
$$E'\Big(\sup_{v\le t}\Big[\int_0^v \int p_{t-s}(y-x)dM(s,y)\Big]^2\Big) < \infty.$$

(b) Let X and u be as above. There is a filtered space $\bar{\Omega}'' = (\Omega'', \mathcal{F}'', \mathcal{F}''_t, P'')$ so that

$$\bar{\Omega} = (\Omega, \mathcal{F}, \mathcal{F}_t, P) \equiv (\Omega' \times \Omega'', \mathcal{F}' \times \mathcal{F}'', (\mathcal{F}'_\cdot \times \mathcal{F}''_\cdot)_{t+}, P' \times P'')$$

carries an $\mathcal{F}_t$-white noise, W, on $\mathbb{R}_+ \times \mathbb{R}$ and $u \circ \Pi$ satisfies $(\mathrm{SPDE})_{X_0}$ on $\bar{\Omega}$, where $\Pi : \Omega \to \Omega'$ is the projection map.
(c) Assume u satisfies $(\mathrm{SPDE})_m$ (d-dimensional) on some $\bar{\Omega} = (\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Then

$$X_t(dx) = \begin{cases} u(t,x)dx & \text{if } t > 0 \\ m(dx) & \text{if } t = 0 \end{cases}$$

defines an $(\mathcal{F}_t)$-SBM(γ) on $\bar{\Omega}$ starting at m.

A proof is given below. Clearly Theorem III.3.8(c) is immediate from (a).

Corollary III.4.3. (a) If $d = 1$, then for any $m \in M_F(\mathbb{R})$ there is a solution to $(\mathrm{SPDE})_m$ and the law of u on $C((0,\infty), C_K(\mathbb{R}))$ is unique.
(b) If $d \geq 2$ and $m \neq 0$, there is no solution to $(\mathrm{SPDE})_m$.
Proof. (a) The existence is included in Theorem III.4.2(b). The Borel subsets of $C_K(\mathbb{R})$ are generated by the coordinate maps. To prove uniqueness in law it therefore suffices to show that if u satisfies $(\mathrm{SPDE})_m$, then

$$(III.4.2) \quad P((u_{t_i}(x_i))_{i \leq n} \in \cdot) \text{ is unique on } \mathcal{B}(\mathbb{R}^n) \text{ for any } 0 < t_i,\ x_i \in \mathbb{R},\ n \in \mathbb{N}.$$

If X is the SBM(γ) in Theorem III.4.2(c), then $u_{t_i}(x_i) = \lim_{\varepsilon \to 0} X_{t_i}(p_\varepsilon(\cdot - x_i))$, and the uniqueness in law of the super-Brownian motion X clearly implies (III.4.2).

(b) If u is a solution to $(\mathrm{SPDE})_m$ for $d \geq 2$, then by Theorem III.4.2(c), $X_t(dx) = u(t,x)dx$ for $t > 0$ and $X_0 = m$ defines a super-Brownian motion starting at m which is absolutely continuous for $t > 0$. This contradicts the a.s. singularity of super-Brownian motion in dimensions greater than 1 (Theorem III.3.8(a), but note that Proposition III.3.1 suffices if $d \geq 3$). ∎

The uniqueness in law for the above stochastic pde does not follow from the standard theory since the square root function is not Lipschitz continuous. For Lipschitz continuous functions of the solution (as opposed to $\sqrt{\gamma u}$) solutions to (SPDE) are unique in law when the initial condition has a nice density (see Chapter 3 of Walsh (1986)). This needs naturally to
Open Problem. Does pathwise uniqueness hold in $(\mathrm{SPDE})_m$? That is, if u, v are solutions of $(\mathrm{SPDE})_m$ with the same white noise W, is it true that $u = v$ a.s.?

Note that the finite-dimensional version of this problem is true by Yamada-Watanabe (1971).

Recently Mytnik (1998) proved uniqueness in law for solutions of $(\mathrm{SPDE})_m$ when the square root in front of the white noise is replaced by a power between 1/2 and 1. His argument may be viewed as an extension of the exponential duality used to prove uniqueness for superprocesses but the dual process is now random. It does apply for slightly more general functions than powers but, as is often the case with duality arguments, the restriction on the functions is severe and artificial. This is one reason for the interest in the above problem as a pathwise uniqueness argument would likely be quite robust.

To prove Theorem III.4.2 we use the following version of Kolmogorov's continuity criterion for two-parameter processes.

Proposition III.4.4. Let $I:(t_0,\infty)\times\mathbb{R}\to\mathbb{R}$ be a process on $(\Omega,\mathcal{F},P)$ such that for some $p>1$, $a,b>2$, for any $T>t_0$, there is a $c=c(T)$ so that

$$E(|I(t',x')-I(t,x)|^p)\le c(T)[|t'-t|^a+|x'-x|^b]\ \forall t,t'\in(t_0,T],\ x,x'\in[-T,T].$$

Then I has a continuous version.
Proof. See Corollary 1.2 of Walsh (1986) where a modulus of continuity of the continuous version is also given.

Lemma III.4.5. (a) If $0\le\delta\le p$, then

$$|p_{t+\varepsilon}(z)-p_t(z)|^p\le(\varepsilon t^{-3/2})^\delta[p_{t+\varepsilon}(z)^{p-\delta}+p_t(z)^{p-\delta}]\quad\forall z\in\mathbb{R},\ t>0,\ \varepsilon\ge 0.$$

(b) If $0<\delta<1/2$, there is a $c(\delta)>0$ so that for any $0\le t\le t'\le T$ and $x,x'\in\mathbb{R}$,

$$\int_0^{t'}\int(p_{t'-s}(y-x')-p_{t-s}(y-x))^2\,dy\,ds\le|x'-x|+c(\delta)T^{1/2-\delta}|t'-t|^\delta.$$

Proof. (a) By the mean value theorem there is a $u\in[t,t+\varepsilon]$ such that

$$\begin{aligned}|p_{t+\varepsilon}(z)-p_t(z)|=\varepsilon\left|\frac{\partial p_u}{\partial u}(z)\right|&=\varepsilon\frac{p_u(z)}{2u}\left|\frac{z^2}{u}-1\right|\\&\le\varepsilon u^{-3/2}\le\varepsilon t^{-3/2},\end{aligned}$$

where a calculus argument has been used to see that

$$\sqrt{u}p_u(z)/2\left|\frac{z^2}{u}-1\right|\le(2\pi)^{-1/2}\left[\sup_{x>1/2}xe^{-x}\vee\frac{1}{2}\right]\le 1.$$

Therefore for $0\le\delta\le p$,

$$\begin{aligned}|p_{t+\varepsilon}(z)-p_t(z)|^p&\le(\varepsilon t^{-3/2})^\delta|p_{t+\varepsilon}(z)-p_t(z)|^{p-\delta}\\&\le(\varepsilon t^{-3/2})^\delta\left[p_{t+\varepsilon}(z)^{p-\delta}+p_t(z)^{p-\delta}\right].\end{aligned}$$

(b) Note that if $0\le t\le t'\le T$, then

$$\begin{aligned}&\int_0^{t'}\int(p_{t'-s}(y-x')-p_{t-s}(y-x))^2\,dy\,ds\\&\quad=\int_t^{t'}\int p_{t'-s}(y-x')^2\,dy\,ds+\int_0^t\int(p_{t'-s}(y-x')-p_{t-s}(y-x))^2\,dy\,ds\end{aligned}$$

$(III.4.3)\equiv I_1+I_2.$

By Chapman-Kolmogorov,

$$(III.4.4)\qquad I_1=\int_t^{t'}p_{2(t'-s)}(0)\,ds=\pi^{-1/2}(t'-t)^{1/2}.$$

If we expand the integrand in I_2, let $\Delta = x' - x$, and use Chapman-Kolmogorov, we get

$$\begin{aligned} I_2 &= \int_0^t p_{2(t'-s)}(0) + p_{2(t-s)}(0) - 2p_{t'-s+t-s}(\Delta)\, ds \\ &= \int_0^t p_{2(t'-s)}(0) - p_{t'-s+t-s}(\Delta)\, ds + \int_0^t p_{2(t-s)}(0) - p_{t'-s+t-s}(\Delta)\, ds \\ &\equiv I_3 + I_4. \end{aligned}$$

Consider I_3 and use (a) with $p = 1$ and $0 < \delta < 1/2$ to see that

$$\begin{aligned} I_3 &= \int_0^t (p_{2(t'-s)}(0) - p_{2(t'-s)}(\Delta)) + (p_{2(t'-s)}(\Delta) - p_{t'-s+t-s}(\Delta))\, ds \\ &\le \frac{1}{2}\int_0^{2t'} p_s(0) - p_s(\Delta)\, ds \\ &\quad + (t'-t)^\delta \int_0^t (t'-s)^{-\delta 3/2} \left[p_{2(t'-s)}(\Delta)^{1-\delta} + p_{t'+t-2s}(\Delta)^{1-\delta}\right]\, ds \\ &\le 2^{-3/2}\pi^{-1/2} \int_0^{2t'} s^{-1/2}[1 - \exp(-\Delta^2/(2s))]\, ds \\ &\quad + (t'-t)^\delta \int_0^t (t-s)^{-\delta 3/2 - (1-\delta)/2} (4\pi)^{-(1-\delta)/2} 2\, ds \\ &\le (4\pi)^{-1}|\Delta| \int_{\Delta^2/2t'}^\infty u^{-3/2}(1 - e^{-u})\, du + 1.1(t'-t)^\delta \int_0^t s^{-1/2-\delta}\, ds \\ &\le (4\pi)^{-1}|\Delta| \int_0^\infty u^{-3/2}(1 \wedge u)\, du + 1.1(1/2-\delta)^{-1} t^{1/2-\delta}(t'-t)^\delta \\ &\le \pi^{-1}|\Delta| + c'(\delta) T^{1/2-\delta}(t'-t)^\delta. \end{aligned}$$

Use this, the analogous bound for I_4, and (III.4.4) in (III.4.3) to conclude that the left-hand side of (III.4.3) is bounded by

$$\pi^{-1/2}|t'-t|^{1/2} + |x'-x| + c''(\delta) T^{1/2-\delta}(t'-t)^\delta \le |x'-x| + c(\delta) T^{1/2-\delta}(t'-t)^\delta. \quad \blacksquare$$

We let $p_\varepsilon^x(y) = p_\varepsilon(y - x)$. To use Proposition III.4.4 we will need the following bound on the moments of X_t.

Lemma III.4.6. If X is as in Theorem III.4.2 (a), then

$$E'(X_t(p_\varepsilon^x)^p) \le p!\gamma^p t^{p/2} \exp(X_0(1)/\gamma t) \quad \forall t, \varepsilon > 0,\ x \in \mathbb{R},\ \text{and } p \in \mathbb{N}.$$

Proof. We apply Lemma III.3.6 with $f = \theta p_\varepsilon^x$, where $\theta = \gamma^{-1} t^{-1/2}$ and $\varepsilon, t > 0$ and $x \in \mathbb{R}$ are fixed. Then

$$\frac{\gamma}{2} G(f,t) \equiv \frac{\gamma}{2} \int_0^t \sup_{x'} P_s f(x')\, ds = \frac{\gamma\theta}{2} \int_0^t p_{s+\varepsilon}(0)\, ds \le 1/2,$$

and so Lemma III.3.6 implies that

$$E'(\exp(\theta X_t(p_\varepsilon^x))) \le \exp(2\theta X_0(p_{t+\varepsilon}^x)) \le \exp(X_0(1)/\gamma t).$$

This shows that for any $p \in \mathbb{N}$,

$$E'(X_t(p_\varepsilon^x)^p) \le p!\theta^{-p}\exp(X_0(1)/\gamma t),$$

as is required. ∎

Proof of Theorem III.4.2. (a) We adapt the argument of Konno and Shiga (1988). From (GFR) (see Exercise II.5.2) we see that for each fixed $\varepsilon, t > 0$ and $x \in \mathbb{R}$,

$$(III.4.5) \qquad X_t(p_\varepsilon^x) = X_0(p_{t+\varepsilon}^x) + \int_0^t \int p_{t+\varepsilon-s}(y-x)dM(s,y) \text{ a.s.}$$

Lemma III.4.5 (a) with $\delta = p = 1$ implies that

$$(III.4.6) \qquad \lim_{\varepsilon\downarrow 0} \sup_{x\in\mathbb{R}, t\ge\eta} |X_0(p_{t+\varepsilon}^x) - X_0(p_t^x)| \le 2\varepsilon\eta^{-3/2}X_0(1) \quad \forall \eta > 0.$$

To take L^2 limits in the stochastic integral in (III.4.5) apply Lemma III.4.5(a) with $p = 2$ and $0 < \delta < 1/2$ to see that

$$\begin{aligned}
&E'\left(\int_0^t \int (p_{t+\varepsilon-s}(y-x) - p_{t-s}(y-x))^2 X_s(dy)\, ds\right) \\
&= \int \left[\int_0^t \int (p_{t+\varepsilon-s}(y-x) - p_{t-s}(y-x))^2 p_s(y-z)\, dy\, ds\right] m(dz) \\
&\le \int \Big[\int_0^t \int (\varepsilon(t-s)^{-3/2})^\delta [p_{t+\varepsilon-s}(y-x)^{2-\delta} + p_{t-s}(y-x)^{2-\delta}] \\
&\qquad\qquad \times p_s(y-z)dy\, ds\Big] m(dz) \\
&\le \varepsilon^\delta \int \Big[\int_0^t \int |t-s|^{-3\delta/2}[(t+\varepsilon-s)^{(\delta-1)/2} p_{(t+\varepsilon-s)/(2-\delta)}(y-x) \\
&\qquad + (t-s)^{(\delta-1)/2} p_{(t-s)/(2-\delta)}(y-x)] p_s(y-z)\, dy\, ds\Big] m(dz) \\
&\le 2\varepsilon^\delta \int \Big[\int_0^t (t-s)^{-1/2-\delta}\Big(p_{(t+\varepsilon-s)/(2-\delta)+s}(x-z) \\
&\qquad\qquad + p_{(t-s)/(2-\delta)+s}(x-z)\Big)\, ds\Big] m(dz) \\
&\le 4(2\pi)^{-1/2} m(1)\varepsilon^\delta \left[\int_0^t (t-s)^{-1/2-\delta} s^{-1/2}\, ds\right] \\
&= m(1)c(\delta)t^{-\delta}\varepsilon^\delta.
\end{aligned}$$

This implies that

(III.4.7)

$$\lim_{\varepsilon\downarrow 0} \sup_{x\in\mathbb{R}, t\ge\eta} E'\left(\left(\int_0^t \int p_{t+\varepsilon-s}(y-x) - p_{t-s}(y-x)dM(s,y)\right)^2\right) = 0 \ \forall \eta > 0,$$

and also shows that
$(III.4.8)$

$$\int_0^u \int p_{t-s}(y-x)dM(s,y), \quad u \le t \ \text{ is a continuous } L^2\text{-bounded martingale.}$$

By (III.4.6) and (III.4.7) we may take L^2 limits in (III.4.5) as $\varepsilon \downarrow 0$ and also choose $\varepsilon_n \downarrow 0$ so that for any $t > 0$ and $x \in \mathbb{R}$,

$$\lim_{n\to\infty} X_t(p^x_{\varepsilon_n}) = X_0(p^x_t) + \int_0^t \int p_{t-s}(y-x)dM(s,y) \text{ a.s. and in } L^2.$$

Therefore if we define $u(t,x) = \liminf_{n\to\infty} X_t(p^x_{\varepsilon_n})$ for all $t > 0,\ x \in \mathbb{R}$, then

$$(III.4.9)\ \ u(t,x) = X_0(p^x_t) + \int_0^t \int p_{t-s}(y-x)dM(s,y) \text{ a.s. for all } t > 0 \text{ and } x \in \mathbb{R}.$$

Also standard differentiation theory shows that for each $t > 0$ with probability 1, $X_t(dx) = u(t,x)dx + X^s_t(dx)$, where X^s_t is a random measure such that $X^s_t \perp dx$. Now (III.4.8) and (III.4.9) imply that

$$E'\left(\int u(t,x)\,dx\right) = \int X_0(p^x_t)\,dx = X_0(1) = E'(X_t(1)).$$

This shows that $E'(X^s_t(1)) = 0$ and so

$$(III.4.10) \qquad X_t(dx) = u(t,x)dx \text{ a.s. for all } t > 0$$

Now fix $t_0 > 0$. Apply (III.4.5) with t replaced by t_0 and ε replaced by $t - t_0 > 0$ to obtain

$$X_{t_0}(p^x_{t-t_0}) = X_0(p^x_t) + \int_0^{t_0} \int p_{t-s}(y-x)dM(s,y) \text{ a.s. } \forall t > t_0,\ x \in \mathbb{R}.$$

This and (III.4.9) show that

$$\begin{aligned} u(t,x) &= X_{t_0}(p^x_{t-t_0}) + \int_{t_0}^t \int p_{t-s}(y-x)dM(s,y) \text{ a.s. } \forall t > t_0,\ x \in \mathbb{R} \\ (III.4.11) \qquad &\equiv X_{t_0}(p^x_{t-t_0}) + I(t,x). \end{aligned}$$

Proposition III.4.4 is now used to obtain a continuous version of I as follows. Let $0 < t_0 < t \le t' \le T$, $x, x' \in \mathbb{R}$, and $p > 1$. Then Burkholder's inequality and

(III.4.10) show that (recall that $p_t(x) = 0$ if $t \le 0$)

$$\begin{aligned}
&E'(|I(t',x') - I(t,x)|^{2p}) \\
&\quad \le c_p E'\left(\left(\int_{t_0}^{t'} \gamma \int (p_{t'-s}(y-x') - p_{t-s}(y-x))^2 u(s,y)\,dy\,ds\right)^p\right) \\
&\quad \le \gamma^p \left(\int_{t_0}^{t'} \int (p_{t'-s}(y-x') - p_{t-s}(y-x))^2\,dy\,ds\right)^{p-1} \\
&\qquad\quad \times \int_{t_0}^{t'} \int (p_{t'-s}(y-x') - p_{t-s}(y-x))^2 E'(u(s,y)^p)\,dy\,ds,
\end{aligned}$$

by Jensen's inequality. Lemma III.4.6 and Fatou's Lemma show that

$$(III.4.12) \qquad E'(u(t,y)^p) \le p!\gamma^p t^{p/2} \exp(X_0(1)/\gamma t) \quad \forall t > 0,\ x \in \mathbb{R},\ p \in \mathbb{N}.$$

This, together with Lemma III.4.5(b) and the previous inequality, show that for any $0 < \delta < 1/2$,

$$\begin{aligned}
E'\left(|I(t',x') - I(t,x)|^{2p}\right) &\le c_p\gamma^p[|x-x'| + c(\delta)T^{1/2-\delta}(t'-t)^\delta]^{p-1} \\
&\quad \times p!T^{p/2}\exp(X_0(1)/\gamma t_0)[|x-x'| + c(\delta)T^{1/2-\delta}(t'-t)^\delta] \\
(III.4.13) \qquad &\le c(p,\gamma,X_0(1),t_0,T)[|x-x'|^p + |t'-t|^{p\delta}].
\end{aligned}$$

Proposition III.4.4 shows there is a continuous version of I on $(t_0,\infty)\times\mathbb{R}$. Dominated Convergence shows that $(t,x) \mapsto X_{t_0}(p^x_{t-t_0})$ is also continuous on $(t_0,\infty)\times\mathbb{R}$ a.s., and so (III.4.11) shows there is continuous version of $u(t,x)$ on $(t_0,\infty)\times\mathbb{R}$ for all $t_0 > 0$ and hence on $(0,\infty)\times\mathbb{R}$. We also denote this continuous version by u. Clearly $u(t,x) \ge 0$ for all $(t,x) \in (0,\infty)\times\mathbb{R}$ a.s. Define a measure-valued process by $\tilde{X}_t(dx) = u(t,x)dx$. Then (III.4.10) shows that

$$(III.4.14) \qquad \tilde{X}_t = X_t \text{ a.s. for each } t > 0.$$

If $\phi \in C_K(\mathbb{R})$, then $t \mapsto \tilde{X}_t(\phi)$ is continuous on $(0,\infty)$ a.s. by Dominated Convergence and the continuity of u. Therefore the weak continuity of X and (III.4.14) imply

$$\tilde{X}_t(\phi) = X_t(\phi)\ \forall t > 0 \text{ a.s.},$$

and hence

$$X_t(dx) = \tilde{X}_t(dx) \equiv u(t,x)dx\ \forall t > 0 \text{ a.s.},$$

where u is jointly continuous. Since $\cup_{t\ge\eta}S(X_t)$ is compact for all $\eta > 0$ a.s. by Corollary III.1.7, and u is uniformly continuous on compact sets, it follows easily that $t \mapsto u(t,\cdot)$ is a continuous function from $(0,\infty)$ to $C_K(\mathbb{R})$. (III.4.1) holds by (III.4.8) and (III.4.9), and so the proof of (a) is complete.

(b) Let $(\Omega'', \mathcal{F}'', \mathcal{F}''_t, P'')$ carry an $(\mathcal{F}''_t)$-white noise W'' on $\mathbb{R}_+ \times R$, and define W on $\bar{\Omega}$ by

$$W_t(A) = \int_0^t \int 1_A(x) 1(u(s,x) > 0)(\gamma u(s,x))^{-1/2} dM(s,x) + \int_0^t \int 1_A(x) 1(u(s,x) = 0) dW''(s,x),$$

for $t > 0$ and $A \in \mathcal{B}_F(\mathbb{R})$. The first stochastic integral is a continuous square integrable $(\mathcal{F}'_t)$-martingale with square function $\int_0^t \int 1_A(x) 1(u(s,x) > 0)\, dx\, ds$ because the integrand is in the space $\mathcal{L}^2$ from Remark II.5.5(b). It is now easy to check that $W_t(A)$ is a continuous $(\mathcal{F}_t)$-martingale on $\bar{\Omega}$ with square function $|A|t$ and so is an $(\mathcal{F}_t)$-Brownian motion with diffusion parameter $|A|$. Clearly if A, B are disjoint in $\mathbb{B}_F(\mathbb{R})$, then $W_t(A \cup B) = W_t(A) + W_t(B)$ a.s. and so W is an $(\mathcal{F}_t)$-white noise on $\mathbb{R}_+ \times \mathbb{R}$. We may extend the stochastic integrals with respect to M and W'' to $\mathbb{P}(\mathcal{F}.) \times \mathcal{B}(\mathbb{R})$-measurable integrands because M and W'' are both orthogonal martingale measures with respect to $\mathcal{F}_t$ (we suppress the projection maps in our notation). It follows easily from the definition of W that if $\psi \in \mathcal{L}^2_{\mathrm{loc}}(W)$, then

$$W_t(\psi) = \int_0^t \int \psi(s,\omega,x) 1(u(s,x) > 0)(\gamma u(s,x))^{-1/2} dM(s,x) + \int_0^t \int \psi(s,\omega,x) dW''(s,x).$$

Therefore if $\phi \in C_b^2(\mathbb{R})$, then $\int_0^t \int \phi(x)(\gamma u(s,x))^{1/2} dW(s,x) = M_t(\phi)$. The martingale problem for $X_t(dx)$ $(= u(t,x)dx$ if $t > 0)$ now shows that u satisfies $(\mathrm{SPDE})_m$.

(c) The fact that $t \mapsto u(t,\cdot)$ is a continuous $C_K(\mathbb{R}^d)$-valued process shows that X_t is a continuous $M_F(\mathbb{R}^d)$-valued process for $t > 0$. As was noted in Remark III.4.1, $(\mathrm{SPDE})_m$ implies that $\lim_{t \to 0+} X_t = m$ a.s. and so $X.$ is a.s. continuous on $[0, \infty)$. If $\phi \in C_b^2(\mathbb{R}^d)$ and $M_t(\phi) = \int_0^t \int \phi(x)(\gamma u(s,x))^{1/2} dW(s,x)$, then $M_t(\phi)$ is an $\mathcal{F}_t$-local martingale satisfying

$$\langle M(\phi) \rangle_t = \int_0^t \int \gamma \phi(x)^2 u(s,x)\, dx\, ds = \int_0^t X_s(\gamma \phi^2)\, ds.$$

Therefore $(\mathrm{SPDE})_m$ and Remark II.5.11 imply that X_t satisfies $(\mathrm{LMP})_{\delta_m}$ and so X is an $(\mathcal{F}_t)$-SBM(γ) starting at m. ∎

5. Polar Sets

Throughout this section X is a $SBM(\gamma)$ under $\mathbb{P}_{X_0}$, $X_0 \neq 0$. Recall the range of X is $\mathcal{R} = \bigcup_{\delta > 0} \overline{\mathcal{R}}([\delta, \infty))$.

Definition. If $A \subset \mathbb{R}^d$ we say $X(\omega)$ charges A iff $X_t(\omega)(A) > 0$ for some $t > 0$, and $X(\omega)$ hits A iff $A \cap \mathcal{R}(\omega) \neq \phi$.

Theorem III.5.1. (a) If $\phi \in b\mathcal{B}(\mathbb{R}^d)$, $X_t(\phi)$ is a.s. continuous on $(0, \infty)$.
(b) $\mathbb{P}_{X_0}$ $(X$ charges $A) > 0 \Leftrightarrow A$ has positive Lebesgue measure.

Proof. (a) See Reimers (1989b) or Perkins (1991), and the Remark below.

(b) Since $X_t(A)$ is a.s. continuous by (a), X charges A iff $X_t(A) > 0$ for some rational $t > 0$. The probability of the latter event is positive iff

$$\begin{aligned}\mathbb{P}_{X_0}(X_t(A)) > 0\ \exists t \in \mathbb{Q}^{>0} &\Leftrightarrow P^{X_0}(B_t \in A) > 0\ \exists t \in \mathbb{Q}^{>0}\\ &\Leftrightarrow A \text{ has positive Lebesgue measure.} \quad \blacksquare\end{aligned}$$

Remark. In Theorem 4 of Perkins (1991) (a) is proved for a large class of Y – DW – superdiffusions whose semigroup satisfies a strong continuity condition. The simple idea of the proof is to first use the Garsia-Rodemich-Rumsey Lemma to obtain an explicit modulus of continuity for $X_t(\phi) - X_0(P_t\phi)$ for $\phi \in C_b(E)$ and then to show this modulus is preserved under $\stackrel{bp}{\to}$ of ϕ. The strong continuity condition on P_t implies

$$\int_0^1 \|P_{t+r}\phi - P_t\phi\| r^{-1}\, dr < \infty\ \forall t > 0,\ \phi \in b\mathcal{E},$$

which is already stronger than the continuity of $t \to P_t\phi(x)$ for each $\phi \in b\mathcal{E}$, $x \in E$. The latter is clearly a necessary condition for a.s. continuity of $X_t(\phi)$ on $(0,\infty)$ (take means). Whether or not this, or even norm continuity of $P_t\phi\ \forall\phi \in b\mathcal{E}$, is also sufficient for continuity of $X_t(\phi)$, $t > 0$ remains open.

The notion of hitting a set will be probabilistically more subtle and more important analytically.

Definition. A Borel set is polar for X iff $\mathbb{P}_{X_0}(X \text{ hits } A) = 0$ for all $X_0 \in M_F(\mathbb{R}^d)$ or equivalently iff $\mathbb{P}_{X_0}(X \text{ hits } A) = 0$ for some non-zero X_0 in $M_F(\mathbb{R}^d)$.

The above equivalence is a consequence of the equivalence of $\mathbb{P}_{X_0}$ and $\mathbb{P}_{X_0'}$ on the field $\bigcup_{\delta>0} \sigma(X_r,\ r \ge \delta)$ for any non-zero finite measures X_0 and X_0' (see Example III.2.3). We would like to find an analytic criterion for polarity. For ordinary Brownian motion Kakutani (1944) did this using Newtonian capacity.

Definition. Let g be a decreasing non-negative continuous function on $(0,\infty)$ with $g(0+) > 0$. If $\mu \in M_F(\mathbb{R}^d)$, $\langle \mu \rangle_g = \int\int g(|x-y|)\mu(dx)\mu(dy)$ is the g-energy of μ. If $A \in \mathcal{B}(\mathbb{R}^d)$, let

$$I(g)(A) = \inf\{\langle \mu \rangle_g : \mu \text{ a probability},\ \mu(A) = 1\}$$

and let the g-capacity at A be $C(g)(A) = I(g)(A)^{-1} \in [0,\infty)$.

Note that the g-capacity of A is positive iff A is large enough to support a probability of finite g-energy.

Notation.

$$g_\beta(r) = \begin{cases} r^{-\beta} & \beta > 0 \\ 1 + (\log 1/r)^+ & \beta = 0 \\ 1 & \beta < 0 \end{cases}$$

If $\phi \in \mathcal{H}$ there is a close connection between sets of zero Hausdorff ϕ-measure and sets of zero ϕ^{-1}-capacity:

(III.5.1) $\phi - m(A) < \infty \Rightarrow C(\phi^{-1})(A) = 0$.

(III.5.2) $C(g_\beta)(A) = 0 \Rightarrow x^\beta(\log 1/x)^{-1-\varepsilon} - m(A) = 0\ \forall\varepsilon > 0, \beta > 0$.

(III.5.3) $C(g_0)(A) = 0 \Rightarrow (\log^+ 1/x)^{-1}(\log^+ \log^+ 1/x)^{-1-\varepsilon} - m(A) = 0\ \forall\varepsilon > 0$.

Moreover these implications are essentially best possible. See Taylor (1961) for a discussion. In particular the capacitory dimension (defined in the obvious way) and Hausdorff dimension coincide.

For $d \geq 2$ the range of a Brownian motion $\{B_t : t \geq 0\}$ is two-dimensional and so should hit sets of dimension greater than $d-2$. Kakutani (1944) showed for $d \geq 2$ and $A \in \mathcal{B}(\mathcal{R}^d)$,

$$P(B_t \in A\ \exists t > 0) > 0 \Leftrightarrow C(g_{d-2})(A) > 0.$$

Recall from Theorem III.3.9 (see also Remark III.3.2) that $\mathcal{R}$ is a 4-dimensional set if $d \geq 4$ and hence should hit sets of dimension greater than $d-4$.

Theorem III.5.2. Let $A \in \mathcal{B}(\mathbb{R}^d)$. A is polar for X iff $C(g_{d-4})(A) = 0$. In particular, points are polar for X iff $d \geq 4$.

Remark III.5.3. The inner regularity of the Choquet capacities

$$A \to \mathbb{P}_{\delta_x}(A \cap \bar{\mathcal{R}}([\delta, \infty)) \neq 0)\ (\delta > 0) \text{ and } A \to C(g_{d-4})(A)$$

(see III.29 of Dellacherie and Meyer (1978)) allows one to consider only $A = K$ compact. The necessity of the zero capacity condition for polarity was proved in Perkins (1990) by a probabilistic inclusion-exclusion argument. The elegant proof given below is due to Le Gall (1999) (Section VI.2). The more delicate sufficiency was proved by Dynkin (1991). His argument proceeded in two steps:

1. $V(x) = -\log \mathbb{P}_{\delta_x}(\mathcal{R} \cap K = \phi)$ is the maximal non-negative solution of $\Delta V = \gamma V^2$ on K^c.
2. The only non-negative solution of $\Delta V = \gamma V^2$ on K^c is $V \equiv 0$ iff $C(g_{d-4})(K) = 0$.

Step 1 uses a probabilistic representation of solutions to the non-linear boundary value problem

$$\Delta V = \gamma V^2 \text{ in D, } V|_{\partial D} = g,$$

where D is a regular domain in $\mathbb{R}^d$ for the classical Dirichlet problem, and g is a bounded, continuous, non-negative function on ∂D. The representation is in terms of the exit measure X_D of X on D. X_D may be constructed as the weak limit of the of the sequence of measures obtained by stopping the branching particles in Section II.4 when they exit D and assigning mass $1/N$ to each exit location.

Step 2 is the analytical characterization of the sets of removable singularities for $\Delta V = \gamma V^2$ due to Baras and Pierre (1984). A self-contained description of both steps may be found in Chapter VI of Le Gall (1999). A proof of a slightly weaker result is given below (Corollary III.5.10).

The proof of the necessity of the zero capacity condition in Theorem III.5.2 will in fact show

Theorem III.5.4. For any $M > 0$ and $X_0 \in M_F(\mathbb{R}^d) - \{0\}$, there is a $c(M, X_0) > 0$ such that

$$\mathbb{P}_{X_0}(X \text{ hits } A) \geq c(M, X_0) C(g_{d-4})(A) \text{ for any Borel subset } A \text{ of } B(0, M).$$

In particular points are not polar if $d \leq 3$.

Proof. We may assume $A = K$ is a compact subset of $B(0, M)$ of positive g_{d-4}-capacity by the inner regularity of $C(g_{d-4})$. Choose a probability ν supported by K so that $\mathcal{E} \equiv \int\int g_{d-4}(x-y) d\nu(x) d\nu(y) < \infty$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuous, non-negative, radially symmetric function such that $\{f > 0\} = B(0,1)$ and

$\int f(y)\,dy = 1$. Define $f_\varepsilon(y) = \varepsilon^{-d} f(\varepsilon^{-1}y)$ and $\phi_\varepsilon(y) = f_\varepsilon * \nu(y) \equiv \int f_\varepsilon(y-z)\nu(dz)$. Note that

$$(III.5.4) \qquad \{\phi_\varepsilon > 0\} \subset K^\varepsilon \equiv \{x : d(x,K) < \varepsilon\}.$$

We will use the following elementary consequence of the Cauchy-Schwarz inequality.

Lemma III.5.5. Assume $Z \geq 0$ has mean μ and variance $\sigma^2 < \infty$. Then

$$P(Z > 0) \geq \mu^2/(\mu^2 + \sigma^2).$$

Proof. $E(Z) = E(Z1(Z>0)) \leq E(Z^2)^{1/2} P(Z>0)^{1/2}$. Rearrange to get a lower bound on $P(Z > 0)$. ∎

Apply the above lemma to $Z = \int_1^2 X_s(\phi_\varepsilon)ds$. If $g_{1,2}(x) = \int_1^2 p_s(x)\,ds$, where p_s is the Brownian transition density, then

$$\begin{aligned}\mathbb{P}_{X_0}\Big(\int_1^2 X_s(\phi_\varepsilon)\,ds\Big) &= \int_1^2 P^{X_0}(\phi_\varepsilon(B_s))\,ds \\ &= \int \Big[\int g_{1,2} * X_0(y) f_\varepsilon(y-z)\,dy\Big]\nu(dz) \\ &\to \int g_{1,2} * X_0(z)\nu(dz) \text{ as } \varepsilon \downarrow 0 \\ &> 0.\end{aligned}$$

The above shows that for $\varepsilon < \varepsilon_0$,

$$(III.5.5) \qquad \mathbb{P}_{X_0}\Big(\int_1^2 X_s(\phi_\varepsilon)\,ds\Big) \geq \frac{1}{2}\inf_{|z|\leq M} g_{1,2} * X_0(z) \equiv C_1(X_0, M).$$

By our second moment formula (see Exercise II.5.2(b))

$$\begin{aligned}\mathrm{Var}\Big(\int_1^2 X_s(\phi_\varepsilon)\,ds\Big) &= \int_1^2\int_1^2 \text{ Cov } (X_s(\phi_\varepsilon), X_t(\phi_\varepsilon))\,dsdt \\ &= 2\gamma\int_1^2 dt\int_1^t ds\int_0^s dr X_0 P_r(P_{s-r}\phi_\varepsilon P_{t-r}\phi_\varepsilon) \\ &\leq 2\gamma\Big[\int_0^{1/2} dr\int_1^2 ds\int_s^2 dt X_0 P_r(P_{s-r}\phi_\varepsilon P_{t-r}\phi_\varepsilon) \\ &\quad + \int_{1/2}^2 dr\int_r^2 ds\int_s^2 dt X_0 P_r(P_{s-r}\phi_\varepsilon P_{t-r}\phi_\varepsilon)\Big] \\ (III.5.6) \qquad &\equiv 2\gamma[I_1 + I_2].\end{aligned}$$

In I_1, $s - r \geq 1/2$, $t - r \geq 1/2$ and so

$$P_{t-r}\phi_\varepsilon \vee P_{s-r}\phi_\varepsilon \leq c\int\int f_\varepsilon(y-z)\,dy d\nu(z) = c.$$

This implies (the value of c may change)

$$(III.5.7) \qquad I_1 \leq cX_0(1).$$

If $G_2\phi_\varepsilon(y) = \int_0^2 P_t\phi_\varepsilon(y)\,dt$, then

$$(III.5.8) \qquad I_2 \le \int_{1/2}^2 X_0 P_r(G_2\phi_\varepsilon^2)\,dr \le cX_0(1)\int G_2\phi_\varepsilon(y)^2\,dy.$$

Lemma III.5.6. (a) $G_2\phi_\varepsilon(y) \le C\int g_{d-2}(y-z)\,d\nu(z)$.
(b) $\int g_{d-2}(z_1-y)g_{d-2}(z_2-y)\,dy \le Cg_{d-4}(z_1-z_2)$.
Proof. (a) Since $\int_0^2 p_t(x)dt \le cg_{d-2}(x)$, we have

$$G_2\phi_\varepsilon(y) \le c\int\int g_{d-2}(y-x)f_\varepsilon(x-z)\,dxd\nu(z).$$

The superharmonicity of g_{d-2} implies that the spherical averages of $g_{d-2}(y-x)f_\varepsilon(x-z)$ over $\{x : |x-z)| = r\}$ are at most $g_{d-2}(y-z)f_\varepsilon(r)$. This and the fact that $\int f_\varepsilon(y)\,dy = 1$ allow us to conclude from the above that

$$G_2\phi_\varepsilon(y) \le c\int g_{d-2}(y-z)d\nu(z).$$

(b) Exercise. One approach is to use $g_{d-2}(x) \le c\int_0^1 p_t(x)\,dt$ and Chapman-Kolmogorov. ∎

Use (a) and (b) in (III.5.8) to see that

$$I_2 \le cX_0(1)\int\int g_{d-4}(z_1-z_2)\,d\nu(z_1)\,d\nu(z_2) = cX_0(1)\mathcal{E}.$$

Now use the above with (III.5.5)–(III.5.8) in Lemma III.5.5 to see that

$$\mathbb{P}_{X_0}\Big(\int_1^2 X_s(\phi_\varepsilon)\,ds > 0\Big) \ge \frac{C_1(X_0,M)^2}{C_1(X_0,M)^2 + cX_0(1)(1+\mathcal{E})} \ge \frac{c(X_0,M)}{\mathcal{E}},$$

where we use $\mathcal{E} \ge c_M > 0$ if $K \subset B(0,M)$ in the last line. Now minimize $\mathcal{E}$ to see that

$$\mathbb{P}_{X_0}\Big(\int_1^2 X_s(\phi_\varepsilon)\,ds > 0\Big) \ge c(X_0,M)C(g_{d-4})(K).$$

This implies

$$\begin{aligned}\mathbb{P}_{X_0}(\overline{\mathcal{R}}([1,2]) \cap K \ne \phi) &= \lim_{\varepsilon\downarrow 0}\mathbb{P}_{X_0}(\overline{\mathcal{R}}([1,2]) \cap K^\varepsilon \ne \phi)\\ &\ge \lim_{\varepsilon\downarrow 0}\mathbb{P}_{X_0}\Big(\int_1^2 X_s(\phi_\varepsilon)\,ds > 0\Big),\end{aligned}$$

the last because $S(\phi_\varepsilon) \subset K^\varepsilon$. The above two inequalities complete the proof. ∎

Upper bounds on hitting probabilities appear to require a greater analytic component. We now obtain precise asymptotics for hitting small balls using the Laplace functional equation (LE) from Section II.5. Recall from Theorem II.5.11 that if $\Delta/2$

denotes the generator of Brownian motion, $f \in C_b(\mathbb{R}^d)_+$, and V_t is the unique solution of

$$(SE)_{0,f} \qquad \frac{\partial V}{\partial t} = \frac{\Delta}{2} V_t - \frac{\gamma}{2} V_t^2 + f, \ V_0 = 0,$$

then

$$(LE) \qquad \mathbb{P}_{X_0}(\exp\{-\int_0^t X_s(f)\,ds\}) = \exp(-X_0(V_t)).$$

Recall also that $f(\varepsilon) \sim g(\varepsilon)$ as $\varepsilon \downarrow 0$ means $\lim_{\varepsilon\downarrow 0} f(\varepsilon)/g(\varepsilon) = 1$.

Theorem III.5.7.(a) If $d \leq 3$, then

$$\mathbb{P}_{X_0}(X \text{ hits } \{x\}) = 1 - \exp\Big\{ - \frac{2(4-d)}{\gamma} \int |y-x|^{-2}\,dX_0(y)\Big\}.$$

(b) There is a $c(d) > 0$ such that if $x \notin S(X_0)$, then as $\varepsilon \downarrow 0$

$$\mathbb{P}_{X_0}(X \text{ hits } B(x,\varepsilon)) \sim \begin{cases} \frac{2}{\gamma} \int |y-x|^{-2} dX_0(y)(\log 1/\varepsilon)^{-1} & \text{if } d = 4 \\ \frac{c(d)}{\gamma} \int |y-x|^{2-d}\,dX_0(y)\varepsilon^{d-4} & \text{if } d > 4. \end{cases}$$

(c) There is a $K_d > 0$ so that if $d(x, S(X_0)) \geq 2\varepsilon_0$, then

$$\mathbb{P}_{X_0}(X \text{ hits } B(x,\varepsilon)) \leq \frac{K_d}{\gamma} X_0(1)\varepsilon_0^{2-d} \begin{cases} (\log 1/\varepsilon)^{-1} & \forall \varepsilon \in (0, \varepsilon_0 \wedge \varepsilon_0^2) & \text{if } d = 4 \\ \varepsilon^{d-4} & \forall \varepsilon \in (0, \varepsilon_0) & \text{if } d > 4. \end{cases}$$

Remark. The constant $c(d)$ is the one arising in Lemma III.5.9 below.

Proof. By translation invariance we may assume $x = 0$. Choose $f \in C_b(\mathbb{R}^d)_+$, radially symmetric so that $\{f > 0\} = B(0,1)$. Let $f_\varepsilon(x) = f(x/\varepsilon)$, and let $u^{\lambda,\varepsilon}(t,x)$ be the unique solution of

$$(SE)_\varepsilon \qquad \frac{\partial u}{\partial t} = \frac{\Delta u}{2} - \frac{\gamma u^2}{2} + \lambda f_\varepsilon, \ u_0 = 0.$$

By scaling, $u^{\lambda,\varepsilon}(t,x) = \varepsilon^{-2} u^{\lambda\varepsilon^4,1}(t\varepsilon^{-2}, x\varepsilon^{-1}) \equiv \varepsilon^{-2} u^{\lambda\varepsilon^4}(t\varepsilon^{-2}, x\varepsilon^{-1})$. We have
$(III.5.9)$

$$\mathbb{P}_{X_0}\Big(\exp\Big\{-\lambda \int_0^t X_s(f_\varepsilon)\,ds\Big\}\Big) = \exp\Big\{-\int \varepsilon^{-2} u^{\lambda\varepsilon^4}(t\varepsilon^{-2}, x\varepsilon^{-1})\,dX_0(x)\Big\}.$$

The left side is decreasing in t and λ and so by taking $X_0 = \delta_x$ we see that $u^\lambda(t,x) \uparrow u(x) = u(|x|)$ $(\leq \infty)$ as $t, \lambda \uparrow \infty$. Take limits in (III.5.9) to get

$$\begin{aligned} \mathbb{P}_{X_0}(X_s(B(0,\varepsilon)) = 0\ \forall s > 0) &= \mathbb{P}_{X_0}\Big(\int_0^\infty X_s(f_\varepsilon)\,ds = 0\Big) \\ &= \lim_{\lambda,t\to\infty} \mathbb{P}_{X_0}\Big(\exp\Big\{-\lambda\int_0^t X_s(f_\varepsilon)\,ds\Big\}\Big). \\ (III.5.10) \qquad &= \exp\{-\int \varepsilon^{-2} u(x\varepsilon^{-1})\,dX_0(x)\}. \end{aligned}$$

The left-hand side increases as $\varepsilon \downarrow 0$ and so (take $X_0 = \delta_x$),

$$(III.5.11) \qquad \varepsilon \to \varepsilon^{-2}u(x/\varepsilon) \text{ decreases as } \varepsilon \downarrow 0, \text{ and in particular the radial function } u(x) \text{ is decreasing in } |x|.$$

By taking $\varepsilon = 1$ and $X_0 = \delta_x$ in (III.5.10), we see that

$$(III.5.12) \qquad u(x) = -\log \mathbb{P}_{\delta_x}(X_s(B(0,1)) = 0 \ \forall s > 0).$$

Suppose $|x| > 1$ and $\mathbb{P}_{\delta_x}(X_s(B(0,1)) = 0 \ \forall s > 0) = 0$. Then by the multiplicative property

$$(III.5.13) \qquad \mathbb{P}_{\frac{1}{n}\delta_x}(X_s(B(0,1)) = 0 \ \ \forall s > 0) = 0 \ \ \forall n \in \mathbb{N}.$$

The extinction probability formula (II.5.12) and Historical Modulus of Continuity (Theorem III.1.3) imply ($\delta(3,\omega)$ is as in the latter result) for some C, $\rho > 0$,

$$\mathbb{Q}_{0,\frac{1}{n}\delta_x}(X_{\frac{1}{n}} = 0, \ \delta(3) > 1/n) \geq e^{-2/\gamma} - \frac{C}{\gamma}\frac{1}{n}\frac{1}{n^\rho} \geq \frac{1}{2}e^{-2/\gamma} \text{ if } n \geq n_0.$$

The above event implies $\mathcal{R} \subset B(x, 3h(1/n))$ and so $X_s(B(0,1)) = 0 \ \forall s > 0$ providing $3h(1/n) < |x| - 1$. This contradicts (III.5.13) and so we have proved $u(x) < \infty$ for all $|x| > 1$, and so is bounded on $\{|x| \geq 1+\varepsilon\}$ $\forall \varepsilon > 0$ by (III.5.11). Letting $\lambda, t \to \infty$ in $(SE)_1$, leads one to believe that u solves

$$(III.5.14) \qquad \Delta u = \gamma u^2 \text{ on } \{|x| > 1\}, \ \lim_{|x|\downarrow 1} u(x) = \infty, \ \lim_{|x|\to\infty} u(x) = 0.$$

Lemma III.5.8. As $\lambda, t \uparrow \infty$, $u^\lambda(t,x) \uparrow u(x)$, where for $|x| > 1$, u is the unique non-negative solution of (III.5.14). Moreover u is C^2 on $\{|x| > 1\}$.

Proof. The mild form of (SE) (recall $(ME)_{0,f}$ from Section II.5) gives (B_t is a Brownian motion under P^x)

$$(III.5.15) \qquad u_t^\lambda(x) = -E^x\Big(\int_0^t \frac{\gamma}{2}u_{t-s}^\lambda(B_s)^2\,ds\Big) + \lambda E^x\Big(\int_0^t f(B_s)\,ds\Big).$$

Let C be an open ball with $\overline{C} \subset \{|x| > 1\}$ and let $T_C = \inf\{t : B_t \notin C\}$. If $x \in C$, the strong Markov property shows that

$$\begin{aligned} u_t^\lambda(x) = &E^x\Big(\int_0^{t\wedge T_C} -\frac{\gamma}{2}u_{t-s}^\lambda(B_s)^2\,ds\Big) \\ &+ E^x\Big(E^{B(T_C)}\Big(\int_0^{(t-T_C)^+} -\frac{\gamma}{2}u_{(t-T_C)^+-s}^\lambda(B_s)^2\,ds\Big) \\ &\quad + \lambda E^{B(T_C)}\Big(\int_0^{(t-T_C)^+} f(B_s)\,ds\Big)\Big), \end{aligned}$$

and so by (III.5.15), with $((t-T_C)^+, B(T_C))$ in place of (t,x),

$$u_t^\lambda(x) + E^x\Big(\int_0^{t\wedge T_C} \frac{\gamma}{2}u_{t-s}^\lambda(B_s)^2\,ds\Big) = E^x\big(u_{(t-T_C)^+}^\lambda(B_{T_C})\big).$$

Now let $t, \lambda \to \infty$ and use Monotone Convergence to see

$$(III.5.16) \qquad u(x) + E^x\Big(\int_0^{T_C} \frac{\gamma}{2} u(B_s)^2\, ds\Big) = E^x(u(B_{T_C}))\ \forall x \in C.$$

The righthand side is harmonic and therefore C^2 on C. The second term on the left is $\frac{\gamma}{2}\int_C g_C(x,y)u(y)^2\, dy$, where g_C is the Green function for C. This is C^2 on C by Theorem 6.6 of Port-Stone (1978). Itô's Lemma and (III.5.16) gives

$$E^x\Big(\int_0^{T_C} \frac{\Delta}{2} u(B_s)\, ds\Big)/E^x(T_C) = \frac{\gamma}{2} E^x\Big(\int_0^{T_C} u(B_s)^2\, ds)/E^x(T_C).$$

Now let $C = B(x, 2^{-n}) \downarrow \{x\}$ to see $\Delta u = \gamma u^2$ on $\{|x| > 1\}$.

Let $\mu(t,x)$ and $\sigma^2(t,x)$ be the mean and variance of $X_t(B(0,1))$ under $\mathbb{P}_{\delta_x}$. Then $\mu(|x|-1, x) \to \mu > 0$ as $|x| \downarrow 1$ and $\sigma^2(|x|-1,x) \le \gamma(|x|-1) \to 0$ as $|x| \downarrow 1$, by our moment formulae (Exercise II.5.2). Therefore Lemma III.5.5 and (III.5.12) show that for $|x| > 1$

$$\begin{aligned} e^{-u(x)} &= \mathbb{P}_{\delta_x}(X_s(B(0,1) = 0\ \forall s > 0) \\ &\le 1 - \mathbb{P}_{\delta_x}(X_{|x|-1}(B(0,1)) > 0) \\ &\le \sigma^2(|x|-1,x)/(\mu(|x|-1,x)^2 + \sigma^2(|x|-1,x)) \\ &\to 0 \text{ as } |x| \downarrow 1. \end{aligned}$$

Therefore $\lim_{|x|\downarrow 1} u(x) = \infty$. (III.5.11) shows that for $|x| \ge 2$,

$$(III.5.17) \qquad u(x) = \frac{4}{|x|^2}\Big(\frac{2}{|x|}\Big)^{-2} u\Big(\frac{2x}{|x|}\Big/\frac{2}{|x|}\Big) \le \frac{4}{|x|^2}u(2) \to 0 \text{ as } |x| \to \infty.$$

It remains only to show uniqueness in (III.5.14). This is an easy application of the classical maximum principle. Let u, v be solutions and $c > 1$. Then $w(x) = c^2 u(cx)$ solves $\Delta w = \gamma w^2$ on $\{|x| > c^{-1}\}$. Therefore

$$\lim_{|x|\downarrow 1} w(x) - v(x) = -\infty, \ \lim_{|x|\to\infty} (w-v)(x) = 0$$

and the usual elementary argument shows that $w - v$ cannot have a positive local maximum. Therefore $w \le v$ and so (let $c \downarrow 1$) $u \le v$. By symmetry $u = v$. ∎

In view of (III.5.10) to obtain the required estimates for Theorem III.5.7 we need to know the asymptotic behavior of $u(x) = u(|x|)$ as $|x| \to \infty$. By the radial symmetry (III.5.14) is an ordinary differential equation in the radial variable and so precise results are known.

Lemma III.5.9. As $r \to \infty$

$$u(r) \sim \begin{cases} \frac{2}{\gamma}(4-d)r^{-2} & d \le 3 \\ \frac{2}{\gamma}r^{-2}/\log r & d = 4 \\ \frac{c(d)}{\gamma}r^{2-d},\ c(d) > 0 & d > 4 \end{cases}.$$

Proof. Iscoe (1988) gives just enough hints to reduce this to a calculus exercise (albeit a lengthy one if $d = 4$) before referring the reader to the differential equations literature. We will carry out this exercise later in this Section to give a self-contained proof. ∎

We are ready to complete the

Proof of Theorem III.5.7. Consider first the proof of (b) and (c) when $d = 4$ (a similar argument works if $d > 4$). From (III.5.10), we have

$$(III.5.18) \qquad \begin{aligned} \mathbb{P}_{X_0}(X \text{ hits } B(0,\varepsilon)) &= \mathbb{P}_{X_0}(X_s(B(0,\varepsilon)) > 0\ \exists s > 0) \\ &= 1 - \exp\{-\varepsilon^{-2} \int u(x/\varepsilon)\, dX_0(x)\}. \end{aligned}$$

Let $2\varepsilon_0 = d(0, S(X_0)) > 0$. If $x \in S(X_0)$ and $0 < \varepsilon < \varepsilon_0 \wedge \varepsilon_0^2$, then the monotonicity of u and Lemma III.5.9 show that

$$(III.5.19) \qquad \begin{aligned} (\log 1/\varepsilon)\varepsilon^{-2} u(x/\varepsilon) &\le (\log 1/\varepsilon)\varepsilon^{-2} u(2\varepsilon_0/\varepsilon) \\ &\le C(\log 1/\varepsilon)\varepsilon^{-2}(\varepsilon_0/\varepsilon)^{-2}(\log(\varepsilon_0/\varepsilon))^{-1} \\ &\le C 2\varepsilon_0^{-2}. \end{aligned}$$

This proves the left side is uniformly bounded on $S(X_0)$ and so (III.5.18), Lemma III.5.9 and Dominated Convergence imply

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} (\log 1/\varepsilon) \mathbb{P}_{X_0}(X \text{ hits } B(0,\varepsilon)) &= \lim_{\varepsilon \downarrow 0} \int (\log 1/\varepsilon)\varepsilon^{-2} u(x/\varepsilon)\, dX_0(x) \\ &= \int \lim_{\varepsilon \downarrow 0} (\log 1/\varepsilon)\varepsilon^{-2} u(x/\varepsilon)\, dX_0(x) \\ &= \frac{2}{\gamma} \int |x|^{-2} dX_0(x). \end{aligned}$$

This proves (b). To prove (c) use (III.5.18) and then (III.5.19) to see that for $\varepsilon < \varepsilon_0 \wedge \varepsilon_0^2$

$$\begin{aligned} \mathbb{P}_{X_0}(X \text{ hits } B(0,\varepsilon)) &\le \varepsilon^{-2} \int u(x/\varepsilon)\, dX_0(x) \\ &\le 2C\varepsilon_0^{-2} X_0(1)(\log 1/\varepsilon)^{-1}. \end{aligned}$$

For (a) we consider 3 cases.

Case 1. $0 \notin S(X_0)$.

Let $2\varepsilon_0 = d(0, S(X_0))$. By (III.5.11), and as in the proof of (b), if $x \in S(X_0)$ and $0 < \varepsilon < \varepsilon_0$,

$$\varepsilon^{-2} u(x/\varepsilon) \le \varepsilon^{-2} u(2\varepsilon_0/\varepsilon) \le \varepsilon_0^{-2} u(2).$$

This allows us to use Dominated Convergence and Lemma III.5.9 to let $\varepsilon \downarrow 0$ in (III.5.10) and conclude

$$(III.5.20) \qquad \begin{aligned} &\mathbb{P}_{X_0}\Big(\cup_{\varepsilon > 0} \{X_s(B(0,\varepsilon)) = 0\ \forall s > 0\} \Big) \\ &\qquad = \exp\Big\{ -\int \frac{2}{\gamma}(4-d)|x|^{-2}\, dX_0(x) \Big\}. \end{aligned}$$

The event on the left hand side clearly implies $0 \notin \mathcal{R}$. Conversely suppose $0 \notin \mathcal{R}$. The historical modulus of continuity and $0 \notin S(X_0)$ imply $0 \notin \overline{\mathcal{R}}([0,\delta])$ for some $\delta > 0$ w.p. 1. Therefore $0 \notin \overline{\mathcal{R}}([0,\infty))$ and so $d(0, \overline{\mathcal{R}}([0,\infty))) > 0$ which implies for some $\varepsilon > 0$ $X_s(B(0,\varepsilon)) = 0$ for all $s \geq 0$. Therefore (III.5.20) is the required equation.

Case 2. $0 \in S(X_0)$, $X_0(\{0\}) = 0$.

For $\eta > 0$, let $X_0 = X_0^{1,\eta} + X_0^{2,\eta}$, where $dX_0^{1,\eta}(x) = 1(|x| > \eta)\, dX_0(x)$. If $\delta > 0$,

$$\begin{aligned}\mathbb{P}_{X_0}(0 \notin \overline{\mathcal{R}}([\delta,\infty))) &= \mathbb{P}_{X_0^{1,\eta}}(0 \notin \overline{\mathcal{R}}([\delta,\infty)))\mathbb{P}_{X_0^{2,\eta}}(0 \notin \bar{\mathcal{R}}([\delta,\infty))) \\ &\geq \mathbb{P}_{X_0^{1,\eta}}(0 \notin \overline{\mathcal{R}})\mathbb{P}_{X_0^{2,\eta}}(X_\delta = 0) \\ &= \exp\Big\{\frac{-2(4-d)}{\gamma}\int 1(|x|>\eta)|x|^{-2}dX_0(x)\Big\}\exp\Big\{\frac{-2X_0^{2,\eta}(1)}{\gamma\delta}\Big\},\end{aligned}$$

where we have applied Case 1 to $X_0^{1,\eta}$ and used the extinction probability formula (II.5.12) in the last line. Let $\eta \downarrow 0$ and then $\delta \downarrow 0$ to get the required lower bound. For the upper bound just use $\mathbb{P}_{X_0}(0 \notin \mathcal{R}) \leq \mathbb{P}_{X_0^{1,\eta}}(0 \notin \mathcal{R})$, apply case 1 to $X_0^{1,\eta}$, and let $\eta \downarrow 0$.

Case 3. $X_0(\{0\}) > 0$.

Note that $\mathbb{P}_{X_0}(X_\delta(\{0\})) = P^{X_0}(B_\delta = 0) = 0$ and so we may use the Markov property and apply the previous cases a.s to X_δ to conclude

$$\begin{aligned}\mathbb{P}_{X_0}(X \text{ misses } \{0\}) &\leq \mathbb{P}_{X_0}(\mathbb{P}_{X_\delta}(X \text{misses } \{0\})) \\ &= \mathbb{P}_{X_0}\Big(\exp\Big\{-\frac{2(4-d)}{\gamma}\int |x|^{-2}dX_\delta(x)\Big\}\Big) \to 0 \text{ as } \delta \downarrow 0\end{aligned}$$

because the weak continuity and Fatou's lemma shows

$$\liminf_{\delta\downarrow 0}\int |x|^{-2}dX_\delta(x) = \infty \text{ a.s.}$$

Therefore X hits $\{0\}$ $\mathbb{P}_{X_0}$-a.s. and the result holds in this final case. ∎

Corollary III.5.10. Let

$$f_\beta(r) = \begin{cases} r^\beta & \text{if } \beta > 0 \\ (\log 1/r)^{-1} & \text{if } \beta = 0\end{cases},$$

and $d \geq 4$. If $A \in \mathcal{B}(\mathbb{R}^d)$ and $f_{d-4} - m(A) = 0$, then A is polar for X. In particular, points are polar for X.

Proof. Assume without loss of generality that A is bounded. Let $A \subset \cup_{i=1}^\infty B(x_i^n, r_i^n)$ where $\lim_{n\to\infty}\sum_{i=1}^\infty f_{d-4}(2r_i^n) = 0$. Choose $x_0 \notin \overline{A}$ and n large so that $d(x_0, \cup_1^\infty B(x_i^n, r_i^n)) = \varepsilon_0 > 0$. Then Theorem III.5.7(c) implies that for n large,

$$\begin{aligned}\mathbb{P}_{\delta_{x_0}}(\mathcal{R}\cap A \neq \phi) &\leq \sum_{i=1}^\infty \mathbb{P}_{\delta_{x_0}}(\mathcal{R}\cap B(x_i^n, r_i^n) \neq \phi) \\ &\leq \frac{K_d}{\gamma}\varepsilon_0^{2-d}\sum_{i=1}^\infty f_{d-4}(r_i^n) \\ &\to 0 \text{ as } n \to \infty. \quad ∎\end{aligned}$$

Remark. (III.5.1) shows that this result is also a consequence of Theorem III.5.2 as the hypothesis on A is stronger than that in Theorem III.5.2. However, in practice it is often easier to verify the Hausdorff measure condition than the capacity condition, and (III.5.2)–(III.5.3) show the conditions are close.

Exercise III.5.1. Let

$$\tilde{\psi}_d(r) = \begin{cases} r^4 \log^+ 1/r & \text{if } d = 4 \\ r^4 & \text{if } d > 4 \end{cases}.$$

Use Theorem III.5.7(c) to show that $\tilde{\psi}_d - m(\overline{\mathcal{R}}([\delta,\infty))) < \infty$ $\forall \delta > 0$ $\mathbb{P}_{X_0}$ – a.s. Conclude that $\mathcal{R}$ is $\mathbb{P}_{X_0}$ a.s. Lebesgue null if $d \geq 4$.

Hint. You may assume without loss of generality that $X_0 = \delta_0$. Why?

Proof of Lemma III.5.9. By considering $v = \gamma u$, we may assume without loss of generality that $\gamma = 1$. Radial symmetry shows that (III.5.14) is equivalent to the ordinary differential equation

$$(III.5.21) \qquad u''(r) + \frac{d-1}{r}u'(r) = u(r)^2 \text{ for } r > 1$$

with the associated boundary conditions, or equivalently

$$(III.5.22) \qquad (r^{d-1}u')' = r^{d-1}u(r)^2 \text{ for } r > 1, \quad \lim_{r\to\infty} u(r) = 0, \quad \lim_{r\to 1+} u(r) = \infty.$$

This shows that $r^{d-1}u'$ is non-decreasing and, as it is non-positive by (III.5.11), we conclude that

$$(III.5.23) \qquad -c_0(d) = \lim_{r\to\infty} r^{d-1}u'(r) \leq 0.$$

If $u(r_0) = 0$ for some $r_0 > 1$, then (III.5.12) implies that $X_1(B(0,1)) = 0$ $\mathbb{P}_{\delta_{r_0}}$-a.s. This contradicts the fact that $\mathbb{P}_{\delta_{r_0}}(X_1(B(0,1)) = P^{r_0}(B_1 \in B(0,1)) > 0$ and we have proved that

$$(III.5.24) \qquad u(r) > 0 \text{ for all } r > 1.$$

Integrate (III.5.22) twice, and use (III.5.23) and $u(\infty) = 0$ to obtain the integral equation

$$(III.5.25) \qquad u(r) = c_0(d)\int_r^\infty t^{1-d}dt + \int_r^\infty t^{1-d}\Big(\int_t^\infty s^{d-1}u(s)^2 ds\Big)dt.$$

This shows that $c_0(d) = 0$ if $d \leq 2$ (or else the first term would be infinite), and for $d \geq 3$ the above gives

$$(III.5.26) \qquad u(r) = \frac{c_0(d)}{d-2}r^{2-d} + \int_r^\infty \Big(\frac{r^{2-d} - s^{2-d}}{d-2}\Big)s^{d-1}u(s)^2 ds.$$

We claim that

$$(III.5.27) \qquad c_0(d) > 0 \text{ iff } d \geq 5.$$

Assume first that $d = 3$ or 4. (III.5.26) implies $u(r) \geq \frac{c_0(d)}{d-2} r^{2-d}$, and therefore for some $c' > 0$,

$$u(r) \geq c' \int_{2r}^{\infty} r^{2-d} s^{d-1} \frac{c_0(d)^2}{(d-2)^2} s^{4-2d} ds \geq c' c_0(d)^2 r^{2-d} \int_{2r}^{\infty} s^{3-d} ds.$$

As the last integral is infinite, this shows $c_0(d)$ must be 0. Assume now that $d \geq 5$ and $c_0(d) = 0$. Then (III.5.26) implies

$$(III.5.28) \qquad u(r) \leq \frac{1}{d-2} r^{2-d} \int_r^{\infty} s^{d-1} u(s)^2 ds \ll r^{2-d} \text{ as } r \to \infty,$$

because (III.5.26) implies the above integral is finite. Use this in (III.5.26) to see there is an $r_0 > 1$ such that $r \geq r_0$ implies

$$\begin{aligned} ru(r) &\leq \frac{1}{d-2} r^{3-d} \int_r^{\infty} s^{d-1} u(s)^2 ds \\ &\leq r^{3-d} \int_r^{\infty} su(s) ds \\ &\leq \int_r^{\infty} su(s) ds, \end{aligned}$$

where the above integral is finite by (III.5.28). Now iterate the above inequality as in Gronwall's lemma to see that $ru(r) = 0$ for $r \geq r_0$, contradicting (III.5.24). This completes the proof of (III.5.27).

Assume $d \geq 5$. Then (III.5.26) implies

$$\begin{aligned} r^{d-2} u(r) &= \frac{c_0(d)}{d-2} + \frac{1}{d-2} \int_r^{\infty} r^{d-2} [r^{2-d} - s^{2-d}] s^{d-1} u(s)^2 ds \\ &\equiv \frac{c_0(d)}{d-2} + h(r). \end{aligned}$$

Clearly $h(r) \leq c' \int_r^{\infty} s^{d-1} u(s)^2 ds \to 0$ as $r \to \infty$ because (III.5.26) implies the above integrals are finite. This proves the required result with $c(d) = \frac{c_0(d)}{d-2}$.

Assume $d \leq 4$ so that (III.5.25) and (III.5.27) give

$$(III.5.29) \qquad u(r) = \int_r^{\infty} t^{1-d} \int_t^{\infty} s^{d-5} (s^2 u(s))^2 ds dt.$$

Recall from (III.5.11) that $r^2 u(r) \downarrow L \geq 0$ as $r \to \infty$. Assume $d \leq 3$ and $L = 0$. If $\varepsilon \in (0,1)$, there is an $r_0(\varepsilon) > 1$ so that $r^2 u(r) \leq \varepsilon$ whenever $r \geq r_0$. Now use (III.5.29) to see that for $r \geq r_0$,

$$r^2 u(r) \leq \varepsilon^2 r^2 \int_r^{\infty} \frac{t^{-3}}{4-d} dt = \frac{\varepsilon^2}{(4-d)2} \leq \frac{\varepsilon}{2}.$$

Iterate the above to see that $u(r) = 0$ for all $r \geq r_0$, which contradicts (III.5.24) and hence proves $L > 0$. The fact that $r^2 u(r) \downarrow L$ and (III.5.29) together imply

$$r^2 \int_r^{\infty} t^{1-d} \int_t^{\infty} s^{d-5} ds dt \, (r^2 u(r))^2 \geq r^2 u(r) \geq L^2 r^2 \int_r^{\infty} t^{1-d} \int_t^{\infty} s^{d-5} ds dt,$$

and therefore

$$\frac{1}{2(4-d)}(r^2u(r))^2 \geq r^2u(r) \geq \frac{L^2}{2(4-d)}.$$

Let $r \to \infty$ to see that $L = \frac{L^2}{2(4-d)}$. This implies that $L = 2(4-d)$ (because $L > 0$) and so the result follows for $d \leq 3$.

It remains to consider the 4-dimensional case which appears to be the most delicate. In this case

$$w(r) \equiv r^2u(r) \downarrow L = 0 \tag{III.5.30}$$

because if L were positive the inner integral in (III.5.29) would be infinite. (III.5.26) shows that (recall $c_0(4) = 0$)

$$\begin{aligned} w(r) &= \frac{1}{2}\int_r^\infty s^3u(s)^2ds - \frac{1}{2}\int_r^\infty su(s)^2dsr^2 \\ &= \frac{1}{2}\Big[\int_r^\infty s^{-1}w(s)^2ds - \int_r^\infty s^{-3}w(s)^2dsr^2\Big] \\ &\equiv \frac{1}{2}\Big[\int_r^\infty s^{-1}w(s)^2ds - g(r)\Big]. \end{aligned} \tag{III.5.31}$$

The monotonicity of w shows that

$$g(r) \leq w(r)^2\int_r^\infty s^{-3}dsr^2 = w(r)^2/2 \downarrow 0 \text{ as } r \to \infty. \tag{III.5.32}$$

Let

$$\begin{aligned} v(r) \equiv (\log r)w(r) &= \frac{1}{2}\log r\int_r^\infty \frac{w(s)^2}{s}ds - \frac{1}{2}(\log r)g(r) \quad \text{by (III.5.31)} \\ &\equiv \frac{1}{2}h(r) - \frac{1}{2}(\log r)g(r) \end{aligned} \tag{III.5.33}$$

Note that by (III.5.32),

$$\frac{\frac{1}{2}(\log r)g(r)}{v(r)} = \frac{1}{2}\frac{g(r)}{w(r)} \leq \frac{1}{4}w(r) \to 0 \text{ as } r \to \infty, \tag{III.5.34}$$

and so

$$\lim_{r\to\infty}\frac{\frac{1}{2}h(r)}{v(r)} = 1. \tag{III.5.35}$$

Now

$$\begin{aligned} h'(r) &= \frac{1}{r}\int_r^\infty \frac{w(s)^2}{s}ds - \frac{(\log r)w(r)^2}{r} \\ &= \frac{1}{r}\Big[2w(r) + g(r) - (\log r)w(r)^2\Big] \quad \text{(by } (III.5.31)) \\ &= \frac{w(r)}{r}\Big[2 + \frac{g(r)}{w(r)} - \frac{1}{2}h(r) + \frac{1}{2}(\log r)g(r)\Big]. \end{aligned}$$

(III.5.34) and (III.5.35) imply $\frac{1}{2}(\log r)g(r) = \varepsilon(r)h(r)$, where $\lim_{r\to\infty}\varepsilon(r) = 0$, and (III.5.32) shows that $d(r) = \frac{g(r)}{w(r)} \to 0$ as $r \to \infty$. We can therefore rewrite the above as

$$(III.5.36) \qquad h'(r) = \frac{w(r)}{r}a(r)[b(r) - h(r)],$$

where $\lim_{r\to\infty} a(r) = \frac{1}{2}$, and $\lim_{r\to\infty} b(r) = 4$.

We claim that $\lim_{r\to\infty} h(r)$ exists in $[0,\infty)$. If $h(r) > 4$ for large enough r this is clear since h is eventually decreasing (by (III.5.36)) and bounded below. In a similar way the claim holds if $h(r) < 4$ for large enough r. Assume therefore that $h(r) \le 4$ for some arbitrarily large r and $h(r) \ge 4$ for some arbitrarily large values of r. We claim that $\lim_{r\to\infty} h(r) = 4$. Let $\varepsilon > 0$ and suppose $\limsup_{r\to\infty} h(r) > 4+\varepsilon$. We may choose $r_n \uparrow \infty$ and $s_n \in (r_n, r_{n+1})$ so that $h(r_n) \ge 4+\varepsilon$ and $h(s_n) \le 4$ and then choose $u_n \in [s_{n-1}, s_n]$ so that h has a local maximum at u_n and $h(u_n) \ge 4+\varepsilon$. This implies $h'(u_n) = 0$ which contradicts (III.5.36) for n sufficiently large. We have proved that $\limsup_{r\to\infty} h(r) \le 4$. A similar argument shows that $\liminf_{r\to\infty} h(r) \ge 4$. In this way the claim is established. This together with (III.5.33) and (III.5.34) shows that

$$(III.5.37) \qquad L = \lim_{r\to\infty} v(r) \text{ exists in } \mathbb{R}_+.$$

An argument similar to that for $d \le 3$, and using (III.5.29), shows that $L > 0$.

We can write (III.5.33) as

$$(III.5.38) \qquad v(r) = \frac{1}{2}\log r \int_r^\infty \frac{v(s)^2}{(\log s)^2 s} ds - \varepsilon(r),$$

where $\lim_{r\to\infty}\varepsilon(r) = 0$ ($\varepsilon(r) \ge 0$) by (III.5.34) and (III.5.37). If $L > \varepsilon > 0$, there is an $r_0 > 1$ such that for $r \ge r_0$,

$$\frac{1}{2}(\log r)(L-\varepsilon)^2 \int_r^\infty (\log s)^{-2}s^{-1}ds - \varepsilon \le v(r) \le \frac{1}{2}(\log r)(L+\varepsilon)^2 \int_r^\infty (\log s)^{-2}s^{-1}ds.$$

Let $r \to \infty$ and then $\varepsilon \downarrow 0$ to see that $L = \frac{1}{2}L^2$ and so $L = 2$. The result for $d = 4$ follows. ∎

We next consider the fixed time analogue of Theorem III.5.7. Let $\frac{\Delta}{2}$ continue to denote the generator of Brownian motion. Recall from Theorem II.5.9 and Example II.2.4(a) that if $\phi \in C_b^2(\mathbb{R}^d)_+$ and $V_t\phi \ge 0$ is the unique solution of

$$(SE)_{\phi,0} \qquad \frac{\partial V}{\partial t} = \frac{\Delta V_t}{2} - \frac{\gamma}{2}V_t^2, \quad V_0 = \phi,$$

then

$$(LE) \qquad \mathbb{P}_{X_0}(\exp(X_t(\phi))) = \exp(-X_0(V_t\phi)).$$

Theorem III.5.11. Let $d \geq 3$. There is a constant C_d such that for all $X_0 \in M_F(\mathbb{R}^d)$, all $t \geq \varepsilon^2 > 0$, and all $x \in \mathbb{R}^d$,

$$\begin{aligned}\mathbb{P}_{X_0}(X_t(B(x,\varepsilon)) > 0) &\leq \frac{C_d}{\gamma}\int p_{\varepsilon^2+t}(y-x)X_0(dy)\varepsilon^{d-2}\\ &\leq \frac{C_d}{\gamma}t^{-d/2}X_0(\mathbb{R}^d)\varepsilon^{d-2}.\end{aligned}$$

Proof. Since $\frac{X_t}{\gamma}$ is a SBM(1) starting at X_0/γ (check the martingale problem as in Exercise II.5.5) we clearly may assume $\gamma = 1$. By translation invariance we may assume $x = 0$. Let $\phi \in C_b^2(\mathbb{R}^d)_+$ be a radially symmetric function such that $\{\phi > 0\} = B(0,1)$, let $\phi_\varepsilon(x) = \phi(x/\varepsilon)$ and let $v^{\lambda,\varepsilon}(t,x) \geq 0$ be the unique solution of $(SE)_{\lambda\phi_\varepsilon,0}$ from Theorem II.5.11. By scaling we have

$$v^{\lambda,\varepsilon}(t,x) = \varepsilon^{-2}v^{\lambda\varepsilon^2,1}(t\varepsilon^{-2},x\varepsilon^{-1}) \equiv \varepsilon^{-2}v^{\lambda\varepsilon^2}(t\varepsilon^{-2},x\varepsilon^{-1}).$$

By (LE),

$$(III.5.39)\quad 1-\mathbb{P}_{X_0}(\exp(-\lambda X_t(\phi_\varepsilon))) = 1-\exp\left(-\int \varepsilon^{-2}v^{\lambda\varepsilon^2}(t\varepsilon^{-2},x\varepsilon^{-1})X_0(dx)\right).$$

The left-hand side is increasing in λ, and so by taking $X_0 = \delta_x$ we see that $v^\lambda(t,x) \uparrow v^\infty(t,x) \leq \infty$. Let $\lambda \to \infty$ in (III.5.39) to conclude that

$$(III.5.40)\quad \mathbb{P}_{X_0}(X_t(B(0,\varepsilon)) > 0) = 1-\exp\left(-\int \varepsilon^{-2}v^\infty(t\varepsilon^{-2},x\varepsilon^{-1})X_0(dx)\right)$$

$$(III.5.41)\quad \leq \varepsilon^{-2}\int v^\infty(t\varepsilon^{-2},x\varepsilon^{-1})X_0(dx).$$

We therefore require a good upper bound on v^∞. A comparison of (III.5.40), with $X_0 = \delta_x,\ \varepsilon = 1$, and the extinction probability (II.5.12) shows that

$$(III.5.42)\quad v^\infty(t,x) \leq \frac{2}{t}.$$

To get a better bound for small t and $|x| > 1$ we let $r > 1$ and suppose

$$(III.5.43)\quad \text{there exist } t_n \downarrow 0 \text{ such that } \sup_{|x|\geq r} v^\infty(t_n,x) \to \infty.$$

Then (III.5.40) implies that $\lim_{n\to\infty}\mathbb{P}_{\delta_{x_n}}(X_{t_n}(B(0,1)) > 0) = 1$ for some $|x_n| \geq r$. Therefore by translation invariance,

$$\begin{aligned}\liminf_{n\to\infty}\mathbb{P}_{\delta_0}(X_{t_n}(B(0,r-1)^c) > 0) &= \liminf_{n\to\infty}\mathbb{P}_{\delta_{x_n}}(X_{t_n}(B(x_n,r-1)^c) > 0)\\ &\geq \liminf_{n\to\infty}\mathbb{P}_{\delta_{x_n}}(X_{t_n}(B(0,1)) > 0) = 1.\end{aligned}$$

On the other hand our historical modulus of continuity (recall Corollary III.1.5) shows the left-hand side of the above is 0. Therefore (III.5.43) must be false. This with (III.5.42) proves that

$$(III.5.44) \qquad M(r) \equiv \sup_{|x|\geq r, t>0} v^\infty(t,x) < \infty \quad \forall r > 1.$$

Proposition II.5.12 gives the Feynman-Kac representation

$$(III.5.45) \qquad v^\lambda(t,x) = E^x\Big(\lambda\phi(B_t)\exp\Big(\frac{-1}{2}\int_0^t v^\lambda(t-s,B_s)\,ds\Big)\Big).$$

Use the strong Markov property at $T_r = \inf\{t : |B_t| = r\}$ to see that (III.5.45) implies that if $|x| \geq r > 1$, then

$$\begin{aligned} v^\lambda(t,x) =& E^x\Big(1(T_r < t)\exp\Big(\frac{-1}{2}\int_0^{T_r} v^\lambda(t-s,B_s)\,ds\Big) \\ & \times E^{B_{T_r}}\Big(\lambda\phi(B_{t-T_r})\exp\Big(\frac{-1}{2}\int_0^{t-T_r} v^\lambda(t-T_r-s,B_s)\,ds\Big)\Big)\Big) \\ \leq& E^x(1(T_r < t)v^\lambda(t-T_r,B_{T_r})). \end{aligned}$$

Let $\lambda \to \infty$ and use Monotone Convergence in the above to see that

$$(III.5.46) \qquad v^\infty(t,x) \leq E^x(1(T_r < t)v^\infty(t-T_r,B_{T_r})), \quad |x| \geq r > 1.$$

If we replace T_r by the deterministic time $t-s$ $(0 \leq s \leq t)$ in the above argument we get

$$(III.5.47) \qquad v^\infty(t,x) \leq P_{t-s}(v^\infty(s,\cdot))(x).$$

Combine (III.5.46) with (III.5.44) to see that for $|x| \geq 7 > r = 2$,

$$\begin{aligned} v^\infty(1,x) &\leq M(2)P^x(T_2 < 1) \\ &\leq 2M(2)P^0(|B_1| > |x| - 2) \text{ (by a } d\text{-dimensional reflection principle)} \\ &\leq c_1\exp(-|x|^2/4), \end{aligned}$$

where we used our bound $|x| > 7$ in the last line. Together with (III.5.42), this gives

$$v^\infty(1,x) \leq c_2 p(2,x) \text{ for all } x \in \mathbb{R}^d,$$

and so (III.5.47) with $s = 1$ implies

$$v^\infty(t,x) \leq c_2 p(t+1,x) \text{ for all } t \geq 1 \text{ and } x \in \mathbb{R}^d.$$

Use this in (III.5.41) to conclude that for $t \geq \varepsilon^2$,

$$\begin{aligned} \mathbb{P}_{X_0}(X_t(B(0,\varepsilon)) > 0) &\leq c_2\varepsilon^{-2}\int p(t\varepsilon^{-2}+1, x\varepsilon^{-1})X_0(dx) \\ &= c_2\varepsilon^{d-2}\int p(t+\varepsilon^2,x)X_0(dx). \end{aligned}$$

This gives the first inequality and the second inequality is then immediate. ∎

A corresponding lower bound is now left as an exercise.

Exercise III.5.2. Use Lemma III.5.5 and our first and second moment formulae (Exercise II.5.2) to prove:
(a) If $d \geq 3$, for any $K \in \mathbb{N}$, $\delta > 0$ $\exists \varepsilon_0(\delta, K) > 0$ and a universal constant $c_d > 0$ so that whenever $\frac{X_0(1)}{\gamma} \leq K$,

$$\mathbb{P}_{X_0}(X_t(B(x,\varepsilon)) > 0) \geq \frac{c_d}{\gamma} \int p_t(y-x) X_0(dy) \varepsilon^{d-2} \quad \forall 0 < \varepsilon < \varepsilon_0(\delta, K), \ t \geq \delta.$$

(b) If $d = 2$, show the conclusion of (a) holds with $(\log 1/\varepsilon)^{-1}$ in place of ε^{d-2} and the additional restriction $t \leq \delta^{-1}$.
Hints: (1) Draw a picture or two to convince yourself that

$$P^y(B_t \in B(x,\varepsilon)) \geq c'_d \varepsilon^d p_t(y-x) \quad \forall t \geq \varepsilon^2.$$

(2) If $B = B(0,\varepsilon)$, then $P_s((P_{t-s}1_B)^2) \leq c''_d\Big(\frac{\varepsilon^d}{(t-s)^{d/2}} \wedge 1\Big) P_t 1_B$.

Remark III.5.12. (a) If $d \geq 3$, then Theorem III.5.11 and the above Exercise give sharp bounds on $\mathbb{P}_{X_0}(X_t(B(x,\varepsilon)) > 0$ as $\varepsilon \downarrow 0$ except for the value of the constants C_d and c_d. Theorem 3.1 of Dawson-Iscoe-Perkins (1989) shows that there is a universal constant $c(d) > 0$ such that

$$(III.5.48) \qquad \lim_{\varepsilon \downarrow 0} \varepsilon^{2-d} \mathbb{P}_{X_0}(X_t(B(x,\varepsilon)) > 0) = c(d) \int p_t(y-x) X_0(dy).$$

If $d = 2$, a companion upper bound to Example III.5.2(b)

$$\mathbb{P}_{X_0}(X_t(B(x,\varepsilon)) > 0) \leq C_2 t^{-1} X_0(1) |\log \varepsilon|^{-1} \quad \forall t \in [\varepsilon, \varepsilon^{-1}] \ \ \forall \varepsilon \in (0, 1/2)$$

is implied by Corollary 3 of Le Gall (1994). A version of (III.5.48) has not been obtained in this more delicate case.

(b) The analogue of Theorem III.5.2 for fixed times is
$(III.5.49)$
$A \cap S(X_t) = \emptyset \ \mathbb{P}_{X_0} - \text{a.s. iff } C(g_{d-2})(A) = 0, \quad \forall A \in \mathcal{B}(\mathbb{R}^d), \ t > 0, \ d \geq 1, \ X_0 \neq 0.$

The reader may easily prove the necessity of the capacity condition for fixed time polarity by means of a straightforward adaptation of the proof of Theorem III.5.4. This result was first proved in Perkins (1990) (see Theorem 6.1). As in Corollary III.5.10, Theorem III.5.11, and Le Gall's companion upper bound if $d = 2$, readily show that for $d \geq 2$, and all $t > 0$ and $A \in \mathcal{B}(\mathbb{R}^d)$,

$$(III.5.50) \qquad f_{d-2} - m(A) = 0 \Rightarrow A \cap S(X_t) = \emptyset \ \mathbb{P}_{X_0} - \text{a.s.} \ .$$

This is of course weaker than the sufficiency of the capacity condition in (III.5.49) which is again due to Dynkin (1992).

Parts (a) and (b) of the following Exercise will be used in the next Section.

Exercise III.5.3. Let X^1, X^2 be independent SBM's with branching rate γ which start at $X^1_0, X^2_0 \in M_F(\mathbb{R}^d) - \{0\}$.
(a) If $d \geq 5$ use Theorem III.5.11 to show that $S(X^1_t) \cap S(X^2_t) = \emptyset$ a.s.

(b) Prove that the conclusion of (a) remains valid if $d = 4$.

Hint. One way to handle this critical case is to first apply Theorem III.3.8 (which has not been proved in these notes) to X^1.

(c) If $d \le 3$ show that $P(S(X_t^1) \cap S(X_t^2) \ne \emptyset) > 0$.
Hint. Use the necessity of the capacity condition in (III.5.49) together with (III.5.2) (the connection between capacity and Hausdorff measure) and Corollary III.3.5 (the latter results are only needed if $d > 1$).

6. Disconnectedness of the Support

Our goal in this Section is to study the disconnectedness properties of the support of super-Brownian motion. The results seem to be scattered in the literature, often with proofs that are only sketched, and we will try to collect them in this section and give a careful proof of the main result (Theorem 6.3). This also gives us the opportunity to advertise an intriguing open problem. The historical clusters of Theorem III.1.1 will be used to disconnect the support and we start by refining that result.

Assume we are in the setting of Section II.8: H is the $(Y, \gamma, 0)$-historical process on the canonical space of paths $\Omega_H[\tau, \infty)$ with law $\mathbb{Q}_{\tau,m}$, and γ is a positive constant. Recall the canonical measures $\{R_{\tau,t}(y, \cdot) : t > \tau, y \in D^\tau\}$ associated with H from (II.8.6) and set

$$P^*_{\tau,t}(y, A) = \frac{R_{\tau,t}(y, A)}{R_{\tau,t}(y, M_F(D) - \{0\})} = \frac{\gamma(t-\tau)}{2} R_{\tau,t}(y, A).$$

From Exercise II.7.2 we may interpret $P^*_{\tau,t}$ as the law of H_t starting from an infinitesimal point mass on y at time τ and conditioned on non-extinction at time t.

If ν is a probability on $M_F^\tau(D)$ we abuse our notation slightly and write $\mathbb{Q}_{\tau,\nu}$ for $\int \mathbb{Q}_{\tau,m}\nu(dm)$ and also adopt a similar convention for the laws $\mathbb{P}_\mu$ (μ a probability on $M_F(E)$) of the corresponding superprocess. If $\tau \le s \le t$, define

$$r_{s,t}(H_t) \in M_F(D^s \times D) \text{ by } r_{s,t}(H_t)(A) = H_t(\{y : (y^s, y) \in A\}).$$

Theorem III.6.1. Let $m \in M_F^\tau(D) - \{0\}$, $t > \tau$ and let Ξ be a Poisson point process on $D^\tau \times \Omega_H[t, \infty)$ with intensity

$$\mu(A \times B) = \int 1_A(y)\mathbb{Q}_{t,P^*_{\tau,t}(y,\cdot)}(B)\frac{2}{\gamma(t-\tau)}m(dy).$$

Then (under $\mathbb{Q}_{\tau,m}$)
(a) $r_{\tau,t}(H_t) \stackrel{\mathcal{D}}{\equiv} \int \delta_y \times \nu_t \Xi(dy, d\nu)$,
(b) $(H_u)_{u \ge t} \stackrel{\mathcal{D}}{\equiv} \Big(\int \nu_u \Xi(dy, d\nu)\Big)_{u \ge t}$.

Proof. (a) Let $A_1, \ldots, A_n$ be a Borel partition of D, let $\phi_i : D \to \mathbb{R}_+$ be bounded Borel maps $i = 1, \ldots, n$, and define $H_t^i(\cdot) = H_t(\cdot \cap \{y^\tau \in A_i\})$ and $f : D^\tau \times D \to \mathbb{R}_+$

by $f(y,w)=\sum_{i=1}^n 1_{A_i}(y)\phi_i(w)$. Then (III.1.3) and (II.8.6)(b) imply

$$
\begin{aligned}
&\mathbb{Q}_{\tau,m}(\exp(-\int f(y^\tau,y)H_t(dy))\\
&\quad=\mathbb{Q}_{\tau,m}\left(\exp\left(-\sum_{i=1}^n H_t^i(\phi_i)\right)\right)\\
&\quad=\prod_{i=1}^n \mathbb{Q}_{\tau,m}(\exp(-H_t^i(\phi_i)))\\
(III.6.1)\qquad&\quad=\exp\left(-\sum_{i=1}^n\int\int 1-e^{-\nu(\phi_i)}\frac{2}{\gamma(t-\tau)}P^*_{\tau,t}(y,d\nu)1_{A_i}(y)dm(y)\right).
\end{aligned}
$$

On the other hand

$$
\begin{aligned}
&E\left(\exp\left(-\int\int f(y,w)\nu_t(dw)\Xi(dy,d\nu)\right)\right)\\
&\quad=E(\exp\left(-\int(1-\exp(-\int f(y,w)\nu_t(dw)))\mu(dy,d\nu)\right)\\
&\quad=\exp\left(-\int\int 1-\exp(-\sum_{i=1}^n 1_{A_i}(y)\nu_t(\phi_i))\frac{2}{\gamma(t-\tau)}P^*_{\tau,t}(y,d\nu_t)m(dy)\right)\\
(III.6.2)\quad&=\exp\left(-\sum_{i=1}^n\int\int 1-e^{-\nu(\phi_i)}\frac{2}{\gamma(t-\tau)}P^*_{\tau,t}(y,d\nu)1_{A_i}(y)dm(y)\right).
\end{aligned}
$$

As (III.6.1) and (III.6.2) are equal, the two random measures in (a) are equal in law by Lemma II.5.9.

(b) Let $\{y_i,H^i):i\le N\}$ be the points of Ξ. Then N is Poisson with mean $\frac{2m(D)}{\gamma(t-\tau)}$, given N, $\{y_i:i\le N\}$ are i.i.d with law $m/m(D)$, and

$$
(III.6.3)\qquad \begin{array}{l}\text{given } \sigma(N,(y_i)_{i\le N}),\ \{H^i:i\le N\} \text{ are independent random}\\ \text{processes with } H^i \text{ having law } \mathbb{Q}_{t,P^*_{\tau,t}(y_i,\cdot)}.\end{array}
$$

Therefore conditional on $\sigma(N,(y_i)_{i\le N},(H_t^i)_{i\le N})$, $\bar{H}_\cdot=\sum_{i=1}^N H^i_\cdot$ is a sum of N independent historical processes. Such a sum will clearly satisfy $(HMP)_{t,\bar{H}_t(\omega)}$ and therefore be a historical process itself (this is the multiplicative property of superprocesses). Therefore the conditional law of $(\bar{H}_u)_{u\ge t}$ given $\sigma(N,(y_i)_{i\le N},(H_t^i)_{i\le N})$ is $\mathbb{Q}_{t,\bar{H}_t}$. Part (a) implies that $\bar{H}_t$ has law $\mathbb{Q}_{\tau,m}(H_t\in\cdot)$. The Markov property of H under $\mathbb{Q}_{\tau,m}$ now shows that $(\bar{H})_{u\ge t}$ has law $\mathbb{Q}_{\tau,m}((H_u)_{u\ge t}\in\cdot)$ which is what we have to prove. ∎

We now reinterpret the above result directly in terms of the historical process H on its canonical space $(\Omega_H[\tau,\infty),\mathbb{Q}_{\tau,m})$. We in fact assume that H is the historical process of a super-Brownian motion with constant branching rate γ, although the result and its proof remain valid for any historical process such that H_s has no atoms $\mathbb{Q}_{\tau,m}$-a.s. for any $s>\tau$.

If $\tau\le s<t$ and $y\in D^s$, let $H_t^{s,y}(\cdot)=H_t(\{w\in\cdot:w^s=y\})$, i.e., $H_t^{s,y}$ is the contribution to H_t from descendants of y at time s.

Corollary III.6.2. Let $m \in M_F^\tau(D)$ and $\tau \le s < t$. Assume either $\tau < s$ or m is non-atomic. Then under $\mathbb{Q}_{\tau,m}$:
(a) Conditional on $\mathcal{F}^H[\tau, s+]$, $S(r_s(H_t))$ is the range of a Poisson point process with intensity $\frac{2H_s(\omega)}{\gamma(t-s)}$.
(b) $H_u = \sum_{y\in S(r_s(H_t))} H_u^{s,y}$ for all $u \ge t$ a.s.
(c) Conditional on $\mathcal{F}^H[\tau, s+] \vee \sigma(S(r_s(H_t)))$, $\{(H_u^{s,y})_{u\ge t} : y \in S(r_s(H_t))\}$ are independent processes and for each $y \in S(r_s(H_t))$, $(H_u^{s,y})_{u\ge t}$ has (conditional) law $\mathbb{Q}_{t,P^*_{s,t}(y,\cdot)}$.
Proof. (a) is included in Theorem III.1.1.
(b) Lemma III.1.2 and the Markov property show that $H_u(S(r_s(H_t))^c)_{u\ge t}$ is a continuous martingale starting at 0 and so is identically 0 a.s. (b) now follows from the definition of $H_u^{s,y}$.
(c) Theorem III.3.4 and Exercise II.8.3 show that H_s is non-atomic a.s (use Theorem III.3.8(c) if $d = 1$). Therefore by the Markov property we may assume without loss of generality that $s = \tau$ and m is non-atomic. We must show that conditional on $S(r_\tau(H_t))$,

$$\{(H_u^{\tau,y})_{u\ge t} : y \in S(r_\tau(H_t))\}$$

are independent, and, for each $y \in S(r_\tau(H_t))$, $(H_u^{\tau,y})_{u\ge t}$ has (conditional) law $\mathbb{Q}_{t,P^*_{\tau,t}(y,\cdot)}$. As this will depend only on the law of $(H_u)_{u\ge t}$, by Theorem III.6.1(b) we may assume that Ξ is as in that result and

$$(III.6.4) \qquad H_u = \int \nu_u \Xi(dy, d\nu) \text{ for all } u \ge t.$$

Let $\{(y_i, H^i) : i \le N\}$ be the points of Ξ (as in the proof of Theorem III.6.1(b)) so that (III.6.4) may be restated as

$$(III.6.5) \qquad H_u = \sum_{i=1}^N H_u^i \text{ for all } u \ge t.$$

Theorem III.6.1(a) implies that $\{y_i : i \le N\} = S(r_\tau(H_t))$ and the fact that m is non-atomic means that all the y_i's are distinct a.s. By (II.8.6)(a), $\nu(y^\tau \ne y_i) = 0$ $P^*_{\tau,t}(y_i, \cdot)$-a.a. ν and so $H_t^i(\{y : y^\tau \ne y_i\}) = 0$ a.s. As in the proof of (b) we may conclude that $H_u^i(\{y : y^\tau \ne y_i\}) = 0$ for all $u \ge t$ a.s. This shows that

$$\begin{aligned} H_u^i(\cdot) &= H_u^i(\cdot, y^\tau = y_i) \text{ for all } u \ge t \text{ a.s.} \\ &= H_u(\cdot, y^\tau = y_i) \text{ for all } u \ge t \text{ a.s. (by (III.6.5) and } y_i \ne y_j \text{ if } i \ne j \text{ a.s.)} \\ &= H_u^{\tau,y_i}(\cdot) \text{ for all } u \ge t \text{ a.s.} \end{aligned}$$

The required result now follows from (III.6.3). ∎

We are ready for the main result of this section (a sketch of this proof was given in Perkins (1995b)). In the rest of this Section we assume $\big(X, (\mathbb{P}_\mu)_{\mu\in M_F(E)}\big)$ is a super-Brownian motion with constant branching rate γ and $\big(H, (\mathbb{Q}_{\tau,m})_{(\tau,m)\in\hat{E}}\big)$ is the corresponding historical process.

Theorem III.6.3. If $d \ge 4$, $S(X_t)$ is totally disconnected $\mathbb{P}_{X_0}$-a.s. for each $t > 0$.

Proof. By Exercise II.5.1 (or II.5.3) we may work with $X_t = \tilde{\Pi}(H)_t$ under $\mathbb{Q}_{0,X_0}$. If $t > 0$ is fixed, $0 \le s < t$ and $y \in S(r_s(H_t))$, let $X_t^{s,y}(A) = H_t^{s,y}(y_t \in A) \equiv H_t^{s,y} \circ \hat{\Pi}_t^{-1}(A)$, where $\hat{\Pi}_t$ is the obvious projection map. Let $\varepsilon_n \downarrow 0$, where $0 < \varepsilon_n < t$. Corollary III.6.2(b) implies

$$(III.6.6) \qquad S(X_t) = \cup_{y \in S(r_{t-\varepsilon_n}(H_t))} S(X_t^{t-\varepsilon_n,y}) \ \forall n \ \text{ a.s.}$$

By Corollary III.6.2(c) and (II.8.7), conditional on $\mathcal{F}^H([0,t-\varepsilon_n)) \vee \sigma(S(r_{t-\varepsilon_n}(H_t)))$, $\{X_t^{t-\varepsilon_n,y} : y \in S(r_{t-\varepsilon_n}(H_t))\}$ are independent and $X_t^{t-\varepsilon_n,y}$ has (conditional) law

$$P^*_{t-\varepsilon_n,t}(y, \nu \circ \hat{\Pi}_t^{-1} \in \cdot) = \frac{R_{\varepsilon_n}(y_{t-\varepsilon_n}, \cdot)}{R_{\varepsilon_n}(y_{t-\varepsilon_n}, 1)}.$$

A Poisson superposition of independent copies of $X_t^{t-\varepsilon_n,y}$ has law $\mathbb{P}_{y_{t-\varepsilon_n}}(X_{\varepsilon_n} \in \cdot)$ by (II.7.11) and so Exercise III.5.3 shows that

$$(III.6.7) \qquad \{S(X_t^{t-\varepsilon_n,y}) : y \in S_{t-\varepsilon_n}(H_t)\} \text{ are disjoint for all } n \text{ a.s.}$$

Let $\delta(3,\omega) > 0$ $\mathbb{Q}_{0,X_0}$-a.s. be as in the Historical Modulus of Continuity (Theorem III.1.3). Then that result and Corollary III.6.2(b) show that

$$S(H_t^{t-\varepsilon_n,y}) \subset S(H_t) \subset K(\delta(3,\omega),3)$$

and so by the definition of $X_t^{t-\varepsilon_n,y}$,

$$(III.6.8) \qquad S(X_t^{t-\varepsilon_n,y}) \subset B(y_{t-\varepsilon_n}, 3h(\varepsilon_n)) \ \text{ if } \varepsilon_n \le \delta(3,\omega).$$

Fix ω outside a null set so that (III.6.6), (III.6.7), and (III.6.8) hold, and $\delta(3,\omega) > 0$. Then $S(X_t)$ can be written as the disjoint union of a finite number of closed sets of arbitrarily small diameter and hence is totally disconnected. ∎

If X^1 and X^2 are independent super-Brownian motions starting from X_0^1 and X_0^2, respectively, then (see Theorem IV.3.2(b) below)

$$(III.6.9) \qquad \text{if } d \ge 6, \ S(X_u^1) \cap S(X_u^2) = \emptyset \ \text{ for all } u > 0 \text{ a.s.}$$

Using this in place of Exercise III.5.3 in the above proof we get a version of the above result which holds for all times simultaneously.

Theorem III.6.4. If $d \ge 6$, then $S(X_t)$ is totally disconnected for all $t > 0$ $\mathbb{P}_{X_0}-$ a.s.

Proof. Our setting and notation is that of the previous argument. Corollary III.6.2 implies that

$$(III.6.10)$$
$$S(X_u) = \cup_{y \in S(r_{(j-1)\varepsilon_n}(H_{j\varepsilon_n}))} S(X_u^{(j-1)\varepsilon_n,y}) \ \ \forall u \in [(j\varepsilon_n, (j+1)\varepsilon_n] \ \forall j,n \in \mathbb{N} \ \text{ a.s.}$$

Corollary III.6.2(c), Exercise II.8.3 and (II.8.7) show that, conditional on

$$\mathcal{F}^H([0,(j-1)\varepsilon_n]) \vee \sigma(S(r_{(j-1)\varepsilon_n}(H_{j\varepsilon_n}))),$$

$\{(X_u^{(j-1)\varepsilon_n,y})_{u\geq j\varepsilon_n} : y \in S(r_{(j-1)\varepsilon_n}(H_{j\varepsilon_n})\}$ are independent and $(X_{u+j\varepsilon_n})_{u\geq 0}$ has law $\mathbb{P}_{\mu_n(y_{(j-1)\varepsilon_n})}$ where $\mu_n(x,\cdot) = R_{\varepsilon_n}(x,\cdot)/R_{\varepsilon_n}(x,1)$. It therefore follows from (III.6.9) that

$$(III.6.11) \qquad \{S(X_u^{(j-1)\varepsilon_n,y}) : y \in S(r_{(j-1)\varepsilon_n}(H_{j\varepsilon_n}))\} \text{ are disjoint}$$
$$\text{for all } u \in (j\varepsilon_n, (j+1)\varepsilon_n]\ \forall j,n \in \mathbb{N} \text{ a.s.}$$

Finally use the Historical Modulus of Continuity and Corollary III.6.2(b), as in the proof of Theorem III.6.3 to see that

$$(III.6.12)\, S(X_u^{(j-1)\varepsilon_n,y}) \subset B(y_{(j-1)\varepsilon_n}, 6h(\varepsilon_n)) \ \text{ forall } u \in [j\varepsilon_n, (j+1)\varepsilon_n]$$
$$\forall j \in \mathbb{N} \text{ and } n \text{ such that } 2\varepsilon_n < \delta(3,\omega).$$

As before (III.6.10)-(III.6.12) show that with probability one, for all $u > 0$, $S(X_u)$ may be written as a finite disjoint union of closed sets of arbitrarily small diameter and so is a.s. totally disconnected. ∎

In one spatial dimension the existence of a jointly continuous density for X (see Section III.3.4) shows that the closed support cannot be totally disconnected for any positive time with probability one. This leaves the

Open Problem. In two or three dimensions, is the support of super-Brownian motion a.s. totally disconnected at a fixed time?

Nothing seems to be known in two dimensions and the only result in this direction for three dimensions is

Theorem 6.5. (Tribe (1991)) Let $\mathrm{Comp}(x)$ denote the connected component of $S(X_t)$ containing x. If $d \geq 3$, then $\mathrm{Comp}(x) = \{x\}$ for X_t-a.a. x $\mathbb{P}_{X_0}$-a.s. for each $t > 0$.

Tribe's result leaves open the possibility that there is a non-trivial connected component in $S(X_t)$ having mass 0. The proof considers the history of a particle x chosen according to X_t and decomposes the support at time t into the cousins which break off from this trajectory in $[t-\varepsilon, t]$ and the rest of the population. He then shows that with positive probability these sets can be separated by an annulus centered at x. By taking a sequence $\varepsilon_n \downarrow 0$ and using a zero-one law he is then able to disconnect x from the rest of the support a.s. The status of Theorem 6.5 in two dimensions remains unresolved.

The critical dimension for Theorem III.6.4, i.e., above which the support is totally disconnected for all positive times, is also not known.

7. The Support Process

In this section we give a brief survey of some of the properties of the set-valued process $S(X_t)$. Let $\mathcal{K}$ be the set of compact subsets of $\mathbb{R}^d$. For non-empty K_1, $K_2 \in \mathcal{K}$, let

$$\rho_1(K_1,K_2) = \sup_{x\in K_1} d(x,K_2) \wedge 1,$$
$$\rho(K_1,K_2) = \rho_1(K_1,K_2) + \rho_1(K_2,K_1),$$

and set $\rho(K,\phi) = 1$ if $K \neq \phi$. ρ is the Hausdorff metric on $\mathcal{K}$ and $(\mathcal{K},\rho)$ is a Polish space (see Dugundji (1966), p. 205,253).

Assume X is SBM(γ) under $\mathbb{P}_{X_0}$ and let $S_t = S(X_t)$, $t \ge 0$. By Corollary III.1.4, $\{S_t : t > 0\}$ takes values in $\mathcal{K}$ a.s. Although the support map $S(\cdot)$ is not continuous on $M_F(\mathbb{R}^d)$, an elementary consequence of the weak continuity of X is that

$$(III.7.1) \qquad \lim_{t\to s} \rho_1(S_s, S_t) = 0 \quad \forall s > 0 \quad \text{a.s.}$$

On the other hand the Historical Modulus of Continuity (see Corollary III.1.5) shows that if $0 < t - s < \delta(\omega, 3)$

$$\rho_1(S_t, S_s) = \sup_{x\in S_t} d(x, S_s) \wedge 1 \le 3h(t-s)$$

and so

$$(III.7.2) \qquad \lim_{t\downarrow s} \rho_1(S_t, S_s) = 0 \quad \forall s > 0 \quad \text{a.s.}$$

(III.7.1) and (III.7.2) show that $\{S_t : t > 0\}$ is a.s. right-continuous in $\mathcal{K}$. The a.s. existence of left limits is immediate from Corollary III.1.5 and the following simple deterministic result (see Lemma 4.1 of Perkins (1990)):

If $f : (0,\infty) \to \mathcal{K}$ is such that $\forall\varepsilon > 0\ \exists\delta > 0$ so that
$0 \le t - u < \delta$ implies $f(u) \subset f(t)^\varepsilon \equiv \{x : d\big(x, f(t)\big) < \varepsilon\}$,
then f possesses left and right limits at all $t > 0$.

(III.7.1) shows that $S_s \subset S_{s-}$ for all $s > 0$ a.s. When an "isolated colony" becomes extinct at time s at location F one expect $F \in S_{s-} - S_s$. These extinction points are the only kind of discontinuities which arise. Theorem 4.6 of Perkins (1990) shows that

$$\mathrm{card}(S_{t-} - S_t) = 0 \quad \text{or} \quad 1 \quad \text{for all} \quad t > 0 \quad \text{a.s.}$$

The nonstandard proof given there may easily be translated into one use the historical process. For $d \ge 3$ the countable space-time locations of these extinction points are dense in the closed graph of X,

$$G_0(X) = \overline{\{(t,x) : x \in S_t, t \ge 0\}} = \{(t,x) : x \in S_{t-}, t > 0\} \cup \{0\} \times S_0$$

(see Theorem 4.8 of Perkins (1990)).

Of course if $S_0 \in \mathcal{K}$, the above arguments show that $\{S_t : t \ge 0\}$ is cadlag in $\mathcal{K}$ a.s. Assume now $S(X_0) \in \mathcal{K}$. Theorem III.3.8 suggests that in 2 or more dimensions the study of the measure-valued process X reduces to the study of the $\mathcal{K}$-valued process $S_t = S(X_t)$, as X_t is uniformly distributed over S_t according to a deterministic Hausdorff measure at least for Lebesgue a.a. t a.s. If

$$F = \{S \in \mathcal{K} : h_d - m(S) < \infty\},$$

then, as one can use finite unions of "rational balls" in the definition of $h_d - m$ on compact sets, it is clear that F is a Borel subset of $\mathcal{K}$ and $\Psi : F \to M_F(\mathbb{R}^d)$, given by $\Psi(S)(A) = h_d - m(S \cap A)$, is Borel. The support mapping $S(\cdot)$ is also a Borel function from $M_F^K(\mathbb{R}^d)$, the set of measures with compact support, to $\mathcal{K}$. Theorem III.3.8 implies that for s, $t>0$ and $A\in\mathcal{B}(F)$

$$\begin{aligned}\mathbb{P}_{X_0}\left(S_{t+s}\in A\mid\mathcal{F}_t^X\right)&=\mathbb{P}_{X_t}(S_s\in A)\\&=\mathbb{P}_{\Psi(S_t)}(S_s\in A)\end{aligned}$$

and so $\{S_t:t>0\}$ is a cadlag F-valued Markov process. This approach however, does not yield the strong Markov property. For this we would need a means of recovering X_t from $S(X_t)$ that is valid for all $t>0$ a.s. and although Theorem III.3.8 (b) comes close in $d\geq 3$, its validity for all $t>0$ remains unresolved. Another approach to this question was initiated by Tribe (1994).

Notation. $d\geq 3\quad X_t^\varepsilon(A)=|S(X_t)^\varepsilon\cap A|\varepsilon^{2-d},\quad A\in\mathcal{B}(\mathbb{R}^d)$. Here $|\cdot|$ is Lebesgue measure.

The a.s. compactness of $S(X_t)$ shows $X_t^\varepsilon\in M_F(\mathbb{R}^d)$ $\forall t>0$ a.s.

Theorem III.7.1. (Perkins (1994)) Assume $d\geq 3$. There is a universal constant $c(d)>0$ such that $\lim\limits_{\varepsilon\downarrow 0}X_t^\varepsilon=c(d)X_t$ $\forall t>0$ $\mathbb{P}_{X_0}$-a.s. In fact if ϕ is a bounded Lebesgue-integrable function and $r<\frac{2}{d+2}$ then $\mathbb{P}_{X_0}$-a.s. there is an $\varepsilon_0(\omega)>0$ so that $\sup\limits_{t\geq\varepsilon^{1/4}}|X_t^\varepsilon(\phi)-c(d)X_t(\phi)|\leq\varepsilon^r$ for $0<\varepsilon<\varepsilon_0$.

Remark. $c(d)$ is the constant given in (III.5.21) below which determines the asymptotic behaviour of $\mathbb{P}_{X_0}\big(X_t(B(x,\varepsilon))>0\big)$ as $\varepsilon\downarrow 0$.

It is now easy to repeat the above reasoning with the above characterization of X_t in place of the Hausdorff measure approach to see that $t\to S(X_t)$ is a Borel strong Markov process with cadlag paths.

Notation. $\Phi_\varepsilon:\mathcal{K}\to M_F(\mathbb{R}^d)$ is given by $\Phi_\varepsilon(S)(A)=|A\cap S^\varepsilon|\varepsilon^{2-d}$.
Define $\Phi:\mathcal{K}\to M_F(\mathbb{R}^d)$ by $\Phi(S)=\begin{cases}\lim\limits_{n\to\infty}\Phi_{1/n}(S) & \text{if it exists}\\ 0 & \text{otherwise}\end{cases}$.
$E=\{S'\in\mathcal{K}:S\big(\Phi(S')\big)=S'\}$; $\Omega_E=D([0,\infty),E)$ with its Borel σ-field $\mathcal{F}^E$, canonical filtration $\mathcal{F}_t^E$, and coordinate maps S_t.

It is easy to check that Φ_ε is Borel and hence so are Φ and E. If $d\geq 3$, Theorem III.7.1 implies $S(X_t)\in E$ $\forall t>0$ $\mathbb{P}_{X_0}$-a.s. and so for $S'\in E$ we may define a probability $\mathbb{Q}_{S'}$ on Ω_E by

$$\mathbb{Q}_{S'}(A)=\mathbb{P}_{\Phi(S')}\big(S(X_\cdot)\in A\big).$$

Corollary III.7.2. Assume $d\geq 3$. $(\Omega_E,\mathcal{F}^E,\mathcal{F}_t^E,S_t,\mathbb{Q}_{S'})$ is a Borel strong Markov process with right-continuous E-valued paths.
Proof. See Theorem 1.4 of Perkins (1994). ∎

Note. At a jump time of $S(X_t)$ the left limit of the support process will not be in E because

$$S(\Phi(S(X)_{t-}))=S(X_t)\neq S(X)_{t-}.$$

Open Problem. Is $S(X_t)$ strong Markov for $d=2$?

The potential difficulty here is that $S(X_t)$ could fold back onto itself on a set of positive X_t measure at an exceptional time $t(\omega)$.

IV. Interactive Drifts

1. Dawson's Girsanov Theorem

Our objective is to study measure-valued diffusions which locally behave like DW-superprocesses, much in the same way as solutions to Itô's stochastic differential equations behave locally like Brownian motions. This means that we want to consider processes in which the branching rate, γ, the spatial generator, A, and the drift, g, all depend on the current state of the system, X_t, or more generally on the past behaviour of the system, $X|_{[0,t]}$. One suspects that these dependencies are listed roughly in decreasing order of difficulty. In this Chapter we present a general result of Dawson which, for a large class of interactive drifts, will give an explicit formula for the Radon-Nikodym derivative of law of the interactive model with respect to that of a driftless DW-superprocess.

We will illustrate these techniques with a stochastic model for two competing populations and hence work in a bivariate setting for most of the time. The models will also illustrate the limitations of the method as the interactions become singular. These singular interactions will be studied in the next Section.

Let E_i, Y_i, A_i, and γ_i, $i = 1, 2$ each be as in Theorem II.5.1 and set $\mathcal{E}_i = \mathcal{B}(E_i)$. Let $\Omega^i_X = C(\mathbb{R}_+, M_F(E_i))$ with its Borel σ-field $\mathcal{F}^{X,i}$ and canonical filtration $\mathcal{F}^{X,i}_t$ and introduce the canonical space for our interacting populations,

$$(\Omega^2, \mathcal{F}^2, \mathcal{F}^2_t) = \Big(\Omega^1_X \times \Omega^2_X, \mathcal{F}^{X,1} \times \mathcal{F}^{X,2}, \big(\mathcal{F}^{X,1}_\cdot \times \mathcal{F}^{X,2}_\cdot\big)_{t+}\Big).$$

The coordinate maps on Ω^2 will be denoted by $X = (X^1, X^2)$ and $\mathcal{P}$ will be the σ-field of $(\mathcal{F}^2_t)$-predictable sets in $\mathbb{R}_+ \times \Omega^2$. For $i = 1, 2$, let $m_i \in M_F(E_i)$, and g_i denote a $\mathcal{P} \times \mathcal{E}_i$-measurable map from $\mathbb{R}_+ \times \Omega^2 \times E_i$ to $\mathbb{R}$. A probability $\mathbb{P}$ on $(\Omega^2, \mathcal{F}^2)$ will satisfy $(MP)^m_g$ iff

$$\forall \phi_i \in \mathcal{D}(A_i) \quad X^i_t(\phi_i) = m_i(\phi_i) + \int_0^t X^i_s(A_i\phi - g_i(s, X, \cdot)\phi_i)ds + M^{i,g_i}_t(\phi_i)$$

defines continuous $(\mathcal{F}^2_t)$ – martingales $M^{i,g_i}_t(\phi)$ $(i = 1, 2)$ under $\mathbb{P}$ such that

$$M^{i,g_i}_0(\phi_i) = 0 \text{ and } \langle M^{i,g_i}(\phi_i), M^{j,g_j}(\phi_j)\rangle_t = \delta_{ij} \int_0^t X^i_s(\gamma_i\phi_i^2)ds.$$

Implicit in (MP) is the fact that $\int_0^t X^i_s(|g_i(s, X, \cdot)|)ds < \infty$ for all $t > 0$ $\mathbb{P}$-a.s. We have inserted a negative sign in front of g_i only because our main example will involve a negative drift.

Example IV.1.1. (Competing Species) Take $E_i = \mathbb{R}^d$, $A_i = \Delta/2$, $\gamma_i \equiv 1$. If p_t denotes the Brownian density, $\varepsilon > 0$ and $\lambda_i \geq 0$, let

$$g^\varepsilon_i(x, \mu) = \lambda_i \int p_\varepsilon(x - y)\mu(dy).$$

Consider two branching particle systems, $X^N = (X^{1,N}, X^{2,N})$, as in Section II.3 with independent spatial (Brownian) motions and independent critical binary branching mechanisms but with one change. At $t = i/N$ a potential parent of a 1-particle located at x_1 dies before it can reproduce with probability $g^\varepsilon_1(x_1, X^{2,N}_{t-})/N$. Similarly a potential parent in the 2-population located at x_2 dies with probability

$g_2^\varepsilon(x_2, X_{t-}^{2,N})/N$ before reaching child-bearing age. This means that the effective branching distribution for the i population is

$$\nu^{i,N}(x_i, X^N) = \delta_0 \frac{1}{2}(1 + g_i^\varepsilon(x_i, X_{t-}^{j,N})/N) + \delta_2 \frac{1}{2}(1 - g_i^\varepsilon(x_i, X_{t-}^{j,N})/N) \ (j \neq i)$$

and so depends on the current state of the population X_{t-}^N as well as the location of the parent. Note that $\int k\nu^{i,N}(x_i, X_{t-}^N)(dk) = 1 - g_i^\varepsilon(x_i, X_{t-}^{j,N})/N \ (j \neq i)$ and so $g_i^\varepsilon(x_i, X_{t-}^N)$ plays the role of g_N in (II.3.1).

The two populations are competing for resources and so a high density of 1's near a 2-particle decreases the likelihood of the successful reproduction of the 2 and a high density of 2's has similar detrimental effect on a 1-particle. λ_i is the susceptibility of the i^{th} population and $\sqrt{\varepsilon}$ is the range of the interaction. The method of Section II.4 will show that if $X_0^i = m_i$ is the initial measure of the i^{th} population, then $\{X^N\}$ is tight in $D(\mathbb{R}_+, M_F(\mathbb{R}^d)^2)$ and all limit points are in Ω^2 and satisfy

$$\begin{aligned} &\forall \phi_i \in \mathcal{D}(\Delta/2) \ X_t^i(\phi_i) = m_i(\phi_i) + \int_0^t X_s^i(\Delta\phi_i/2)ds \\ &(CS)_m^{\varepsilon,\lambda} \qquad - \int_0^t \int g_i^\varepsilon(x_i, X_s^j)\phi_i(x_i)X_s^i(dx_i)ds + M_t^i(\phi_i) \quad (i = 1, 2, \ j \neq i), \end{aligned}$$

where $M_t^i(\phi_i)$ are continuous $(\mathcal{F}_t^2)$ – martingales such that

$$M_0^i = 0 \text{ and } \langle M^i(\phi_i), M^j(\phi_j)\rangle_t = \delta_{ij} \int_0^t X_s^i(\phi_i^2)ds.$$

The only technical point concerns the uniform (in N) bound required on $E(X_t^{i,N}(\phi))$ in the analogue of Lemma II.3.3. However, it is easy to couple X^N with branching particle systems with $\lambda_i = 0$, $Z^{i,N}$, $i = 1, 2$, (ignore the interactive killing) so that $X^{i,N} \leq Z^{i,N}$ and so the required bound is immediate from the $\lambda_i = 0$ case. Clearly $(CS)_m^{\varepsilon,\lambda}$ is a special case of $(MP)_g^m$ with

$$g_i(s, X, x) = g_i^\varepsilon(x, X_s^j), \quad j \neq i.$$

First consider $(MP)_g^m$ in what should be a trivial case: $g_i(s, X, x) = g_i^0(x)$ for some $g_i^0 \in C_b(E_i)$. We let $(MP)_{g^0}^m$ denote this martingale problem. Let $\mathbb{P}_{m_i}^{i,g_i^0}$ be the law of the (A_i, γ_i, g_i^0)-DW-superprocess starting at m_i. If $g_i^0 \equiv \theta_i$ is constant, write $\mathbb{P}_{m_i}^{i,\theta_i}$ for this law and write $\mathbb{P}_{m_i}^i$ for $\mathbb{P}_{m_i}^{i,0}$. Clearly $\mathbb{P}_{m_1}^{1,g_1^0} \times \mathbb{P}_{m_2}^{2,g_2^0}$ satisfies $(MP)_{g^0}^m$ but it remains to show that it is the only solution. It is easy to extend the Laplace function equation approach in Section II.5 (see Exercise IV.1.1 below) but another approach is to use the following result which has a number of other interesting applications.

Theorem IV.1.2. (Predictable Representation Property). Let $\mathbb{P}_m$ be the law of the (Y, γ, g)-DW-superprocess starting at m on the canonical space of $M_F(E)$-

valued paths $(\Omega_X, \mathcal{F}_X, \mathcal{F}_t^X)$. If $V \in L^2(\mathcal{F}_X, \mathbb{P}_m)$, there is an f in

$$\mathcal{L}^2 = \{f : \mathbb{R}_+ \times \Omega_X \times E \to \mathbb{R} : f \text{ is } \mathcal{P}(\mathcal{F}_t^X) \times \mathcal{E} - \text{measurable and}$$
$$\mathbb{E}\Big(\int_0^t \int f(s,X,x)^2 \gamma(x) X_s(dx)ds\Big) < \infty \ \forall t > 0\}$$

such that

$$V = \mathbb{P}_m(V) + \int_0^\infty \int f(s,X,x)dM(s,x).$$

Proof. Let N_t be a square integrable $(\mathcal{F}_t^X)$-martingale under $\mathbb{P}_m$. As the martingale problem $(MP)_m$ for the superprocess X is well-posed, we see from Theorem 2 and Proposition 2 of Jacod (1977) that for each $n \in \mathbb{N}$ there is a finite set of functions, $\phi_n^1, \dots \phi_n^{N(n)} \in \mathcal{D}(A)$, and a finite set of $(\mathcal{F}_t^X)$-predictable processes, $h_n^1, \dots, h_n^{N(n)}$ such that $f_n(s,X,x) = \sum_i h_n^i(s,X)\phi_n^i(x) \in \mathcal{L}^2$ and

$$N_t = \mathbb{P}_m(N_0) + \lim_{n\to\infty} \int_0^t \int f_n(s,X,x)dM(s,x)$$

in $L^2(\Omega_X, \mathcal{F}_X, \mathbb{P}_m)$ for each $t \geq 0$. Hence for each such t,

$$\begin{aligned}
&\lim_{n,n'\to\infty} \mathbb{P}_m\Big(\int_0^t \int [f_n(s,X,x) - f_{n'}(s,X,x)]^2 \gamma(x) X_s(dx)ds\Big) \\
&\quad = \lim_{n,n'\to\infty} \mathbb{P}_m\Big(\Big[\int_0^t \int f_n(s,X,x)dM(s,x) - \int_0^t \int f_{n'}(s,X,x)dM(s,x)\Big]^2\Big) \\
&\quad = 0.
\end{aligned}$$

The completeness of $\mathcal{L}^2$ shows that there is an f in $\mathcal{L}^2$ so that

$$\lim_{n\to\infty} \int_0^t \int f_n(s,X,x)dM(s,x) = \int_0^t \int f(s,X,x)dM(s,x)$$

in $L^2(\Omega_X, \mathcal{F}_X, \mathbb{P}_m)$. This shows that any square integrable $(\mathcal{F}_t^X)$-martingale under $\mathbb{P}_m$ is a constant plus the stochastic integral of a process in $\mathcal{L}^2$ with respect to the martingale measure M. This is of course equivalent to the stated result. ∎

Corollary IV.1.3. $\mathbb{P}_{m_1}^{1,g_1^0} \times \mathbb{P}_{m_2}^{2,g_2^0}$ is the unique solution of $(MP)_{g^0}^m$.

Proof. Let $\mathbb{P}$ be any solution of $(MP)_{g^0}^m$. By the uniqueness of the martingale problem for the DW-superprocess (Theorem II.5.1) we know that $\mathbb{P}(X^i \in \cdot) = \mathbb{P}_{m_i}^{i,g_i^0}(\cdot)$. If ϕ_i is a bounded measurable function on Ω_X^i then by the above predictable representation property

$$\phi_i(X^i) = \mathbb{P}_{m_i}^{i,g_i^0}(\phi_i) + \int_0^\infty \int f_i(s,X_i,x)dM^i(s,x) \ \mathbb{P} - a.s., \ i = 1,2.$$

(Note that the martingale measure arising in the martingale problem for X_i alone agrees with the martingale measure in $(MP)_{g^0}^m$ by the usual bootstrapping argument starting with simple functions.) The orthogonality of M^1 and M^2 implies that

$$\mathbb{P}(\phi_1(X_1)\phi_2(X_2)) = \mathbb{P}^{1,g_1^0}_{m_1}(\phi_1)\mathbb{P}^{2,g_2^0}_{m_2}(\phi_2). \quad \blacksquare$$

Exercise IV.1.1. Let $V_t^i\phi_i$ be the unique solution of

$$\frac{\partial V_t^i}{\partial t} = A_i V_t^i - \frac{\gamma_i (V_t^i)^2}{2} - g_i^0 V_t^i \quad V_0^i = \phi_i, \quad \phi_i \in \mathcal{D}(A_i).$$

Let ν be a probability on $M_F(E_1) \times M_F(E_2)$ and define $(LMP)^\nu_{g^0}$ in the obvious manner (ν is the law of X_0 and the martingale terms are now local martingales in general). Show that any solution $\mathbb{P}$ of $(LMP)^\nu_{g^0}$ satisfies

$$\mathbb{P}\big(\exp\{-X_t^1(\phi_1) - X_t^2(\phi_2\}\big) = \int \exp\{-X_0^1(V_t^1\phi_1) - X_0^2(V_t^2\phi_2)\}d\nu(X_0).$$

Conclude that

$$\mathbb{P}(X_t \in \cdot) = \int \mathbb{P}^{1,g_1^0}_{m_1}(X_t^1 \in \cdot) \times \mathbb{P}^{2,g_2^0}_{m_2}(X_t^2 \in \cdot)d\nu(m),$$

and then convince yourself that the appropriate version of Theorem II.5.6 shows that $(LMP)^\nu_{g^0}$ is well-posed.

Consider now a more general martingale problem than $(MP)^m_g$ on a general filtered space $\bar{\Omega}' = (\Omega', \mathcal{F}', \mathcal{F}'_t, \mathbb{P}')$. If $m_i \in M_F(E_i)$, $i = 1, 2$, a pair of stochastic processes $(X^1, X^2) \in \Omega^2$ satisfies $(MP)^m_{C,D}$ iff

$$\forall \phi_i \in \mathcal{D}(A_i) \quad X_t^i(\phi_i) = m_i(\phi_i) + \int X_s^i(A_i\phi_i)ds - C_t^i(\phi_i) + D_t^i(\phi_i) + M_t^i(\phi_i),$$

where $M_t^i(\phi_i)$ is a continuous $(\mathcal{F}'_t)$ – martingale such that

$\langle M^i(\phi_i), M^j(\phi_j)\rangle_t = \delta_{ij} \int_0^t X_s^i(\gamma_i\phi_i^2)ds, M_0^i(\phi_i) = 0,$ and C^i, D^i are continuous, non-decreasing, adapted $M_F(E_i)$ – valued processes, starting at 0.

If $\bar{\Omega}'$ is as above, introduce

$$\Omega'' = \Omega' \times \Omega^2,\ \mathcal{F}'' = \mathcal{F}' \times \mathcal{F}^2,\ \mathcal{F}''_t = (\mathcal{F}'_\cdot \times \mathcal{F}^2_\cdot)_{t+},\ \bar{\Omega}'' = (\Omega'', \mathcal{F}'', \mathcal{F}''_t),$$

let $\omega'' = (\omega', \tilde{X}^1, \tilde{X}^2)$ denote points in Ω'' and let $\Pi : \Omega'' \to \Omega'$ be the projection map.

Proposition IV.1.4. (Domination Principle) Assume X satisfies $(MP)^m_{C,D}$ on $\bar{\Omega}'$ and for some $\theta_i \in C_b(E_i)_+$,

$$(DOM) \quad (D_t^i - D_s^i)(\cdot) \le \int_s^t X_r^i(\theta_i 1(\cdot))dr \text{ (as measures on } E_i) \ \forall s < t,\ i = 1, 2.$$

There is a probability $\mathbb{P}$ on $(\Omega'', \mathcal{F}'')$ and processes $(Z^1, Z^2) \in \Omega^2$ such that
(a) If $W \in b\mathcal{F}'$, then $\mathbb{P}(W \circ \Pi | \mathcal{F}''_t) = \mathbb{P}'(W|\mathcal{F}'_t) \circ \Pi \ \ \mathbb{P} - a.s.$

(b) $X \circ \Pi$ satisfies $(MP)^m_{C\circ\Pi,D\circ\Pi}$ on $\bar{\Omega}''$.
(c) $Z^1,\ Z^2$ are independent, Z^i is an $(\mathcal{F}''_t) - (Y^i, \gamma_i, \theta_i)$-DW superprocess starting at m_i, and $Z^i_t \ge X^i_t \circ \Pi$ on Ω'' $\forall t \ge 0,\ i = 1, 2$.

Remark IV.1.5. Clearly (a) implies that $W = (X, D, C, M)$ on $\bar{\Omega}'$ and $W \circ \Pi$ on $\bar{\Omega}''$ have the same law. More significantly they have the same adapted distribution in the sense of Hoover and Keisler (1984). This means that all random variables obtained from W, respectively $W \circ \Pi$, by the operations of compositions with bounded continuous functions and taking conditional expectation with respect to $\mathcal{F}''_t$, respectively $\mathcal{F}'_t$, have the same laws. Therefore in studying (X, C, D, M) on $\bar{\Omega}'$ we may just as well study $(X, C, D, M) \circ \Pi$ on $\bar{\Omega}''$ and hence may effectively assume (X^1, X^2) is dominated by a pair of independent DW-superprocesses as above. We will do this in what follows without further ado.

Sketch of Proof. The proof of Theorem 5.1 in Barlow, Evans and Perkins (1991) goes through with only minor changes. We sketch the main ideas.
Step 1. DW-superprocesses with immigration.
Assume

$$\mu_i \in M^i_{LF} = \{\mu : \mu \text{ is a measure on } \mathbb{R}_+ \times E_i,\ \mu([0,T] \times E_i) < \infty\ \forall T > 0,\\ \mu(\{t\} \times E_i) = 0\ \forall t \ge 0\},$$

and $\tau \ge 0$. Consider the following martingale problem, denoted $(MP)^i_{\tau,m_i,\mu}$, for a DW-superprocess with immigration μ on some $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$:

$$\forall \phi \in \mathcal{D}(A_i) \quad X_t(\phi) = m_i(\phi) + \int_\tau^t \int \phi(x) d\mu(r, x) + \int_\tau^t X_s(A_i^{\theta_i}\phi) ds + M_t(\phi),$$

$t \ge \tau$, where $M_t(\phi)$, $t \ge \tau$ is a continuous $(\mathcal{F}_t)$-martingale such that

$$M_\tau(\phi) = 0, \text{ and } \langle M(\phi)\rangle_t = \int_\tau^t X_s(\gamma_i \phi^2) ds.$$

Then $(MP)^i_{\tau,m_i,\mu}$ is well-posed and the law $\mathbb{P}^i_{\tau,m_i,\mu}$ of any solution on Ω_X satisfies

$$\mathbb{P}^i_{\tau,m_i,\mu}(\exp(-X_t(\phi)) = \exp\Big\{ - m_i(V^i_{t-\tau}\phi) - \int_\tau^t \int V^i_{t-s}\phi(x) d\mu(s, x)\Big\},$$

where $V^i_t\phi$ is as in Exercise IV.1.1. Moreover $(\Omega_X, \mathcal{F}_X, \mathcal{F}^X_t, X_t, \mathbb{P}^i_{\tau,m_i,\mu})$ is an inhomogeneous Borel strong Markov process and $(\tau, m_i, \mu) \to \mathbb{P}^i_{\tau,m_i,\mu}$ is Borel measurable. The existence of a solution may be seen by approximating μ by a sequence of measures each supported by $\{t_0, \dots, t_m\} \times E_i$ for some finite set of points, and taking the weak limit through an appropriate sequence of the corresponding DW-superprocesses. For any solution to $(MP)^i$, the formula for the Laplace functional and other properties stated above may then be derived just as in Section II.5. Note that the required measurability is clear from the Laplace functional equation, the Markov property and a monotone class argument. (Alternatively, the existence of a unique Markov process satisfying this Laplace functional equation is a special case of Theorem 1.1 of Dynkin and the corresponding martingale problem may then be derived as in Fitzsimmons (1988,1989).)

Step 2. Definition of $\mathbb{P}$.

Set $\mathbb{Q}^i_\mu = \mathbb{P}^i_{0,0,\mu}$ and define

$$F^i_t(\cdot) = \int_0^t X^i_s(\theta^i 1(\cdot))ds - D^i_t(\cdot) + C^i_t(\cdot).$$

Then $F^i(\omega') \in M^i_{LF}$ $\mathbb{P}'$-a.s. and we can define $\mathbb{P}$ on Ω'' by

$$\mathbb{P}(A \times B_1 \times B_2) = \int_{\Omega'} 1_A(\omega')\mathbb{Q}^1_{F^1(\omega')}(B_1)\mathbb{Q}^2_{F^2(\omega')}(B_2)d\mathbb{P}(\omega').$$

This means that under $\mathbb{P}$, conditional on ω', $\tilde{X}^1$ and $\tilde{X}^2$ are independent, and $\tilde{X}^i$ is a $(Y_i, \gamma_i, \theta_i)$-DW-superprocess with immigration F_i. Define $Z^i_t(\omega', \tilde{X}) = X^i_t(\omega') + \tilde{X}^i_t$. For example if θ_i and D_i are both 0, then we can think of $\tilde{X}^i$ as keeping track of the "ghost particles" (and their descendants) killed off by C^i in the X^i population. When it is added to X^i one should get an ordinary DW-superprocess. (a) is a simple consequence of this definition and (b) is then immediate. To prove (c) we show Z satisfies the martingale problem on $\bar{\Omega}''$ corresponding to $(MP)^m_\theta$ and then use Corollary IV.1.3. This is a straightforward calculation (see Theorem 5.1 in Barlow-Evan-Perkins(1991)). The fact that Z^i dominates X^i is obvious. ∎

We now state and prove a bivariate version of Dawson's Girsanov Theorem for interactive drifts (Dawson (1978). The version given here is taken from Evans-Perkins (1994).

Theorem IV.1.6. Assume $\gamma_i(x) > 0$ for all x in E_i, $i = 1, 2$ and $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$-a.s.,

$$(IV.1.1) \qquad \sum_{i=1}^2 \int_0^t \int \frac{g_i(s, X, x)^2}{\gamma_i(x)} X^i_s(dx)ds < \infty, \ \forall t > 0,$$

so that we can define a continuous local martingale under $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$ by

$$R^g_t = \exp\Big\{ \sum_{i=1}^2 \int_0^t \frac{-g_i(s, X, x)}{\gamma_i(x)} dM^{i,0}(s, x) - \frac{1}{2}\int_0^t \int \frac{g_i(s, X, x)^2}{\gamma_i(x)} X^i_s(dx)ds \Big\}.$$

(a) If $\mathbb{P}$ satisfies $(MP)^m_g$ and (IV.1.1) holds $\mathbb{P}$-a.s., then

$$(IV.1.2) \qquad \frac{d\mathbb{P}}{d\mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2}}\Big|_{\mathcal{F}^2_t} = R^g_t,$$

and in particular there is at most one law $\mathbb{P}$ satisfying $(MP)^m_g$ such that (IV.1.1) holds $\mathbb{P}$-a.s.

(b) If $|g_i|^2/\gamma_i(x)$ and $|g_i|$ are uniformly bounded for $i = 1, 2$ then R^g_t is an $(\mathcal{F}^2_t)$-martingale under $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$ and (IV.1.2) defines the unique law $\mathbb{P}$ which satisfies $(MP)^m_g$.

(c) If $X_\cdot \leq X'_\cdot$ (pointwise inequality of measures) implies

$$(IV.1.3) \qquad -\theta\sqrt{\gamma_i(x)} \leq g_i(t, X, x) \leq g_i(t, X', x), \ i = 1, 2 \ \text{ for all } (t, x),$$

for some constant $\theta \geq 0$, then the conclusion of (b) holds.

Proof. (a) Let

$$T_n = \inf\{t : \sum_{i=1}^{2} \int_0^t \Big[\int \Big(\frac{g_i(s,X,x)^2}{\gamma_i(x)} + 1\Big) X_s^i(ds) + 1\Big] ds \geq n\} \ (\leq n).$$

Assume $\mathbb{P}$ satisfies $(MP)_g^m$, (IV.1.1) holds $\mathbb{P}$-a.s., and define

$$\begin{aligned}\tilde{R}^g_{t\wedge T_n} = \exp\Big\{ \sum_{i=1}^{2} \int_0^{t\wedge T_n} & \int \frac{g_i(s,X,x)}{\gamma_i(x)} M^{i,g_i}(ds,dx) \\ & - \frac{1}{2}\int_0^{t\wedge T_n} \int \frac{g_i(s,X,x)^2}{\gamma_i(x)} X_s^i(dx)ds\Big\}.\end{aligned}$$

Then $\tilde{R}^g_{\cdot\wedge T_n}$ is a uniformly integrable $(\mathcal{F}_t^2)$-martingale under $\mathbb{P}$ (e.g., by Theorem III.5.3 of Ikeda-Watanabe (1981)) and so $d\mathbb{Q}_n = \tilde{R}^g_{T_n} d\mathbb{P}$ defines a probability on $(\Omega^2, \mathcal{F}_t^2)$. If $\overset{m}{=}$ denotes equality up to local martingales and $\phi_i \in C_b^2(E_i)$, then integration by parts shows that under $\mathbb{P}$,

$$\begin{aligned} M^{i,0}_{t\wedge T_n}(\phi)\tilde{R}^g_{t\wedge T_n} &= \Big[M^{i,g_i}_{t\wedge T_n}(\phi_i) - \int_0^{t\wedge T_n}\int g_i(s,X,x)\phi_i(x)X_s^i(dx)ds\Big] \\ &\quad\times \Big[1 + \int_0^{t\wedge T_n}\int \tilde{R}^g_s \frac{g_i(s,X,x)}{\gamma_i(x)} M^{i,g_i}(ds,dx)\Big] \\ &\overset{m}{=} -\int_0^{t\wedge T_n}\int \tilde{R}^g_s g_i(s,X,x)\phi_i(x)X_s^i(dx)ds \\ &\quad + \int_0^{t\wedge T_n} \tilde{R}^g_s d\langle M^{i,g_i}(\phi_i), M^{i,g_i}(g_i/\gamma_i)\rangle_s \\ &= 0. \end{aligned}$$

Therefore under $\mathbb{Q}_n$, $M^{i,0}_{t\wedge T_n}$ is an $(\mathcal{F}_t^2)$-local martingale. As $\mathbb{Q}_n \ll \mathbb{P}$ and quadratic variation is a path property, we also have

$$\langle M^{i,0}_{\cdot\wedge T_n}(\phi_i), M^{j,0}_{\cdot\wedge T_n}(\phi_j)\rangle_t = \delta_{ij}\int_0^{t\wedge T_n} X_s^i(\gamma_i\phi_i^2)ds \quad \forall t \geq 0 \ \ \mathbb{Q}_n - a.s.$$

which is uniformly bounded and hence shows $M^{i,0}_{t\wedge T_n}(\phi_i)$ is a $\mathbb{Q}_n$-martingale. Let $\tilde{\mathbb{Q}}_n$ denote the unique law on $(\Omega^2, \mathcal{F}^2)$ such that $\tilde{\mathbb{Q}}_n\big|_{\mathcal{F}^2_{T_n}} = \mathbb{Q}_n\big|_{\mathcal{F}^2_{T_n}}$ and the conditional law of $X_{T_n+\cdot}$ given $\mathcal{F}^2_{T_n}$ is $\mathbb{P}^1_{X^1_{T_n}} \times \mathbb{P}^2_{X^2_{T_n}}$. Then $\tilde{\mathbb{Q}}_n$ satisfies $(MP)_0^m$ and so, by Corollary IV.1.3, $\tilde{\mathbb{Q}}_n = \mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2}$. Therefore (IV.1.1) implies

$$\mathbb{Q}_n(T_n < t) = \mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2}(T_n < t) \to 0 \text{ as } n \to \infty.$$

Since (IV.1.1) holds $\mathbb{P}$-a.s., $\tilde{R}^g_t$ is an $(\mathcal{F}_t^2)$-local martingale under $\mathbb{P}$ and

$$\begin{aligned} \mathbb{P}(\tilde{R}^g_t) &\geq \mathbb{P}(\tilde{R}^g_{t\wedge T_n} 1(T_n \geq t)) \\ &= \mathbb{P}(\tilde{R}^g_{t\wedge T_n}) - \mathbb{P}(\tilde{R}^g_{t\wedge T_n} 1(T_n < t)) \end{aligned} \tag{IV.1.4}$$

$$= 1 - \mathbb{Q}_n(T_n < t) \to 1 \text{ as } n \to \infty.$$

Therefore $\tilde{R}^g_t$ is a a $\mathbb{P}$-martingale and we may define a unique law, $\mathbb{Q}$, on $(\Omega^2, \mathcal{F}^2)$ by $d\mathbb{Q}\big|_{\mathcal{F}^2_t} = \tilde{R}^g_t d\mathbb{P}\big|_{\mathcal{F}^2_t}$ for all $t > 0$. Now repeat the above argument, but without the T_n's, to see that $\mathbb{Q} = \mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2}$. Note here that it suffices to show $M^{i,0}(\phi_i)$ are local martingales as the proof of Corollary IV.1.3 shows the corresponding local martingale problem is well-posed. Therefore

$$d\mathbb{P}\big|_{\mathcal{F}^2_t} = \big(\tilde{R}^g_t\big)^{-1} d(\mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2})\big|_{\mathcal{F}^2_t} = R^g_t d(\mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2})\big|_{\mathcal{F}^2_t} \quad \forall t > 0.$$

(b) Uniqueness is immediate from (a). Let T_n be as in (a), let

$$g^n(s, X, x) = 1(s \le T_n)(g_1(s, X, x), g_2(s, X, x)),$$

and define a probability on $(\Omega^2, \mathcal{F}^2)$ by $d\mathbb{Q}_n = R^g_{T_n} d(\mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2})$. Now argue just as in the proof of (a) to see that $\mathbb{Q}_n$ solves $(MP)^m_{g^n}$. This martingale problem shows that

$$\begin{aligned}\mathbb{Q}_n(X^i_t(1)) &= m_i(1) + \mathbb{Q}_n\Big(\int_0^{t\wedge T_n} X^i_s(g_i(s, X, \cdot))ds\Big)\\ &\le m_i(1) + c\mathbb{Q}_n\Big(\int_0^{t\wedge T_n} X^i_s(1)ds\Big).\end{aligned}$$

The righthand side is finite by the definition of T_n and hence so is the lefthand side. A Gronwall argument now shows that $\mathbb{Q}_n(X^i_t(1)) \le m_i(1)e^{ct}$ and therefore

$$\begin{aligned}&\mathbb{Q}_n\Big(\sum_{i=1}^2 \int_0^t \int \Big[\frac{g_i(s, X, x)^2}{\gamma_i(x)} + 1\Big]\, X^i_s(dx) + 1\, ds\Big)\\ &\qquad \le (c^2 + 1)(m_1(1) + m_2(1))e^{ct}t + 2t \equiv K(t).\end{aligned}$$

This shows that $\mathbb{Q}_n(T_n < t) \le K(t)/n \to 0$ as $n \to \infty$. Argue exactly as in (IV.1.4) to see that $\mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2}(R^g_t) = 1$ and therefore R^g_t is a martingale under this product measure. A simple stochastic calculus argument as in the proof of (a) shows that (IV.1.2) does define a solution of $(MP)^m_g$. Note that, as for $\mathbb{Q}_n$, one sees that $\mathbb{P}(X^i_t(1)) \le m_i(1)e^{ct}$ and so $M^{i,g_i}(\phi)$ is a martingale (and not just a local martingale) because its square function is integrable.

(c) Define T_n, g^n and $\mathbb{Q}_n$ as in the proof of (b). As before, $\mathbb{Q}_n$ satisfies $(MP)^m_{g^n}$. The upper bound on $-g_i$ allows us to apply Proposition IV.1.4 and define processes $Z^i \ge X^i$, $i = 1, 2$ on the same probability space such that (X^1, X^2) has law $\mathbb{Q}_n$ and (Z^1, Z^2) has law $\mathbb{P}^{1,\theta\sqrt{\gamma_1}}_{m_1} \times \mathbb{P}^{2,\theta\sqrt{\gamma_2}}_{m_2}$. The conditions on g_i show that

$$\begin{aligned}\int_0^t \int \Big[\frac{g_i(s, X, x)^2}{\gamma_i(x)} + 1\Big] X^i_s(dx)ds &\le \int_0^t \int \Big[\theta^2 + \frac{g_i^+(s, X, x)^2}{\gamma_i(x)} + 1\Big] X^i_s(dx)ds\\ &\le \int_0^t \int \Big[\theta^2 + \frac{g_i^+(s, Z, x)^2}{\gamma_i(x)} + 1\Big] Z^i_s(dx)ds \qquad (IV.1.5)\\ &\le (\theta^2 + 1) \int_0^t \int \Big[\frac{g_i(s, Z, x)^2}{\gamma_i(x)} + 1\Big] Z^i_s(dx)ds.\end{aligned}$$

This implies that

$$(IV.1.6) \qquad \mathbb{Q}_n(T_n < t) \le \mathbb{P}^{1,\theta\sqrt{\gamma_1}}_{m_1} \times \mathbb{P}^{2,\theta\sqrt{\gamma_2}}_{m_2}(T_{n/(\theta^2+1)} < t).$$

Now (IV.1.1) and the fact that $\mathbb{P}^{1,\theta\sqrt{\gamma_1}}_{m_1} \times \mathbb{P}^{2,\theta\sqrt{\gamma_2}}_{m_2} \ll \mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2}$ on $\mathcal{F}^2_t$ (from (b)) show that (IV.1.1) holds $\mathbb{P}^{1,\theta\sqrt{\gamma_1}}_{m_1} \times \mathbb{P}^{2,\theta\sqrt{\gamma_2}}_{m_2}$-a.s. and therefore the expression on the righthand side of (IV.1.6) approaches 0 as $n \to \infty$. Therefore the same is true for the lefthand side of (IV.1.6) and we can argue as in (IV.1.4) to see that R^g_t is an $(\mathcal{F}^2_t)$-martingale under $\mathbb{P}^1_{m_1} \times \mathbb{P}^2_{m_2}$. A simple stochastic calculus argument, as in (a), shows that (IV.1.2) defines a law $\mathbb{P}$ which satisfies $(MP)^m_g$. Note that initially one gets that the martingale terms in $(MP)^m_g$ are local martingales. As in (b) they are martingales because a simple Gronwall argument using the upper bound on $-g^i$ (use the above stopping times T_n and Fatou's Lemma) shows that $\mathbb{P}(X^i_t(1)) \le m_i(1)e^{ct}$.

For uniqueness assume $\mathbb{P}$ satisfies $(MP)^m_g$. As above we may use Proposition IV.1.4 to define processes X and Z on a common probability space such that $X^i \le Z^i$ $i = 1, 2$, X has law $\mathbb{P}$, and Z has law $\mathbb{P}^{1,\theta\sqrt{\gamma_1}}_{m_1} \times \mathbb{P}^{2,\theta\sqrt{\gamma_2}}_{m_2}$. Recall that we saw that (IV.1.1) holds a.s. with respect to this latter law and so the calculation in (IV.1.5) shows that it holds $\mathbb{P}$-a.s. as well. The uniqueness is therefore consequence of (a). ∎

Remark IV.1.7. (a) Simply take $g_2 = 0$ in the above to get the usual univariate form of Dawson's Girsanov theorem.

(b) In Theorem 2.3 (b) of Evans-Perkins (1994) this result is stated without the monotonicity part of (IV.1.3). This is false as one can easily see by taking $g_1 = 1/X^1_s(1)$ and noting that the total mass of the solution of $(MP)^m_g$ (if it existed) could now become negative because of the constant negative drift. Fortunately all the applications given there are valid because (IV.1.3) holds in each of them.

(c) If $-g_i \le c$ $i = 1, 2$ for some constant c, then $(MP)^m_g$ is equivalent to $(LMP)^m_g$, i.e., $(MP)^m_g$ but now $M^{i,g_i}_t(\phi_i)$ need only be a continuous local martingale. To see this, assume $\mathbb{P}$ satisfies $(LMP)^m_g$ and let $T^i_n = \inf\{t : X^i_t(1) \ge n\}$ $(n > m_i(1))$. Then $M^{i,g_i}_{t\wedge T^i_n}(1)$ is a square integrable martingale because $\langle M^{i,g_i}(1)\rangle_{t\wedge T^i_n} \le \|\gamma_i\|_\infty nt$. We have

$$X^i_t(1) \le m_i(1) + c\int_0^t X^i_s(1)\,ds + M^{i,g_i}_t(1).$$

Take mean values in the above inequality at time $t \wedge T^i_n$ to see that

$$E(X^i_{t\wedge T^i_n}(1)) \le m_i(1) + c\int_0^t E(X^i_{s\wedge T^i_n}(1))\,ds,$$

and so $E(X^i_{t\wedge T^i_n}(1)) \le m_i(1)e^{ct}$. By Fatou's Lemma this implies $E(X^i_t(1)) \le m_i(1)e^{ct}$. Therefore for each $\phi_i \in \mathcal{D}(\frac{\Delta}{2})$, $M^{i,g_i}_t(\phi_i)$ is an L^2-martingale since its square function is integrable.

As a first application of Theorem IV.1.6 we return to

Example IV.1.1. Recall that $(CS)_m^{\varepsilon,\lambda}$ was a special case of $(MP)_g^m$ with $\gamma_i \equiv 1$ and

$$g_i(s,X,x_i) = \lambda_i \int p_\varepsilon(x_i - x_j) X_s^j(dx_j) \ (j \neq i).$$

Clearly the monotonicity condition (IV.1.3) holds with $\theta = 0$ and (IV.1.1) is clear because $g_i(s,X,x_i) \le \lambda_i \varepsilon^{-d/2} X_s^j(1)$. Part (c) of the above theorem therefore shows that the unique solution of $(CS)_m^{\varepsilon,\lambda}$ is $\mathbb{P}_m^\varepsilon$, where, if $\mathbb{P}_m$ is the law of SBM ($\gamma \equiv 1$), then

$$\begin{aligned}\frac{d\mathbb{P}_m^\varepsilon}{d\big(\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}\big)}\Big|_{\mathcal{F}_t^2} = \exp\Big\{ \sum_{i=1}^2 \Big[-\lambda_i \int_0^t \int\int p_\varepsilon(x_i - x_{3-i}) X_s^{3-i}(dx_{3-i}) dM^{i,0}(s,x_i)\\ -\frac{1}{2}\lambda_i^2 \int_0^t \int \Big[\int p_\varepsilon(x_i - x_{3-i}) X_s^{3-i}(dx_{3-i})\Big]^2 X_s^i(dx_i)ds\Big]\Big\}.\end{aligned} \tag{IV.1.7}$$

(IV.1.7) defines a collection of laws $\{\mathbb{P}_m^\varepsilon : m \in M_F(\mathbb{R}^d)\}$ on $(\Omega^2, \mathcal{F}^2)$. If ν is a probability on $M_F(\mathbb{R}^d)^2$ and $\mathbb{P}$ satisfies $(CS)_\nu^{\varepsilon,\lambda}$, that is the analogue of (CS) but with $\mathcal{L}(X_0) = \nu$ and $M_t^i(\phi_i)$ now a local martingale, then one easily sees that the regular conditional probability of X given X_0 satisfies $(CS)_{X_0}^{\varepsilon,\lambda}$ for ν-a.a. X_0. Therefore this conditional law is $\mathbb{P}_{X_0}^\varepsilon$ ν-a.s. and one can argue as in Theorem II.5.6 to see that $(\Omega^2, \mathcal{F}^2, \mathcal{F}_t^2, X_t, \mathbb{P}_m^\varepsilon)$ is a Borel strong Markov process. The Borel measurability is in fact clear from (IV.1.7).

Exercise IV.1.2 Assume Y is a Feller process on a locally compact separable metric space E with strongly continuous semigroup and fix $\gamma > 0$. Let $V_s(\omega) = \omega_s$ be the coordinate maps on $\Omega_V = C(\mathbb{R}_+, M_1(E))$. For each $V_0 \in M_(E)$ there is a unique law $\tilde{\mathbb{P}}_{V_0}$ on Ω_V such that under $\tilde{\mathbb{P}}_{V_0}$

$$\forall \phi \in \mathcal{D}(A) \quad V_t(\phi) = V_0(\phi) + \int_0^t V_s(A\phi)\, ds + M_t(\phi), \text{ where } M(\phi) \text{ is a continuous}$$

$$(\mathcal{F}_t^V)\text{-martingale such that } M_0(\phi) = 0 \text{ and } \langle M(\phi)\rangle_t = \gamma \int_0^t V_s(\phi^2) - V_s(\phi)^2\, ds.$$

$\tilde{\mathbb{P}}_{V_0}$ is the law of the Fleming-Viot process with mutation operator A (see Section 10.4 of Ethier-Kurtz (1986)).

For $c \ge 0$ and $m \in M_1(E)$ consider the following martingale problem for a law $\mathbb{P}$ on Ω_X:

$$\forall \phi \in \mathcal{D}(A) \quad X_t(\phi) = m(\phi) + \int_0^t X_s(A\phi) + c(1 - X_s(1))X_s(\phi)\, ds + M_t^c(\phi), \text{where}$$

$$M^c(\phi) \text{ is an } \mathcal{F}_t^X\text{-martingale such that } M_0^c(\phi) = 0 \text{ and } \langle M^c(\phi)\rangle_t = \gamma \int_0^t X_s(\phi^2)\, ds.$$

(a) Show there is a unique law $\mathbb{P}^c$ satisfying this martingale problem and find $\frac{d\mathbb{P}^c}{d\mathbb{P}^0}|_{\mathcal{F}_t^X}$.
(b) Show that for any $T, \varepsilon > 0$, $\lim_{c\to\infty} \mathbb{P}^c(\sup_{t\le T} |X_t(1) - 1| > \varepsilon) = 0$.

Hint. This is an exercise in one-dimensional diffusion theory–here is one approach. By a time change it suffices to show the required convergence for

$$W_t = 1 + \sqrt{\gamma}B_t + \int_0^t c(1 - W_s)ds.$$

Itô's Lemma implies that for any integer $p \geq 2$,

$$(W_t-1)^p+cp\int_0^t (W_s-1)^p ds = p\sqrt{\gamma}\int_0^t (W_s-1)^{p-1}dB_s+\frac{p(p-1)\gamma}{2}\int_0^t (W_s-1)^{p-2}ds.$$

Use induction and the above to conclude that for each even $p \geq 2$, $\lim_{c\to\infty} E(\int_0^t (W_s - 1)^p ds) = 0$. Now note that the left side of the above display is a nonnegative submartingale. Take $p = 4$ and use a maximal inequality.
(c) Define $S = \inf\{t : X_t(1) \leq 1/2\}$ and $Z_t(\cdot) = \frac{X_{t\wedge S}(\cdot)}{X_{t\wedge S}(1)} \in M_1(E)$. If $\phi \in \mathcal{D}(A)$, prove that $Z_t(\phi) = m(\phi) + \int_0^{t\wedge S} Z_s(A\phi)\, ds + N_t^c(\phi)$, where $N^c(\phi)$ is an $(\mathcal{F}_t^X)$-martingale under $\mathbb{P}^c$ starting at 0 and satisfying

$$\langle N^c(\phi)\rangle_t = \gamma\int_0^{t\wedge S} (Z_s(\phi^2) - Z_s(\phi)^2)X_s(1)^{-1}\, ds.$$

Show this implies $\lim_{c\to\infty} \mathbb{P}^c(Z_t(\phi)) = m(P_t\phi)$.
(d) Show that $\mathbb{P}^c(Z \in \cdot)\overset{w}{\Rightarrow}\tilde{\mathbb{P}}_m$ on Ω_V as $c \to \infty$ and conclude from (b) that $\mathbb{P}^c\overset{w}{\Rightarrow}\tilde{\mathbb{P}}_m$ on Ω_X (we may consider $\tilde{\mathbb{P}}_m$ as a law on Ω_X because $\Omega_V \in \mathcal{F}_X$).
Hint. Use Theorem II.4.1 to show that $\{\mathbb{P}^{c_n}(Z \in \cdot)\}$ is tight for any $c_n \uparrow \infty$. One approach to the compact containment is as follows:

Let d be a bounded metric on $E \cup \{\infty\}$, the one-point compactification of E, let $h_p(x) = e^{-pd(x,\infty)}$ and $g_p(x) = \int_0^1 P_s h_p(x)\, ds$. Then $Ag_p(x) = P_1h_p(x) - h_p(x)$ and (c) gives

$$\sup_{t\leq T} Z_t(g_p) \leq m(g_p) + \sup_{t\leq T} |N_t^c(g_p)| + \int_0^T Z_s(P_1h_p)ds.$$

Now use the first moment result in (c) and a square function inequality to conclude that

$$\lim_{p\to\infty}\limsup_{n\to\infty} \mathbb{P}^{c_n}\Big(\sup_{t\leq T} Z_t(g_p)\Big) = 0.$$

2. A Singular Competing Species Model–Dimension One

Consider $(CS)_m^{\varepsilon,\lambda}$ as the interaction range $\sqrt{\varepsilon} \downarrow 0$. In this limiting regime it is only the local density of the "2-population" at x that has an adverse effect on the "1-population" at x and conversely. It would seem simplest to first study this limiting model in the one-dimensional case where according to the results of Section III.4 we can expect these densities to exist. Throughout this Section $\mathbb{P}_m$ is the law of SBM ($\gamma \equiv 1$) and we continue to use the notation from the last Section with $E_i = \mathbb{R}$, $i = 1, 2$.

Define a Borel map $U : M_F(\mathbb{R}) \times \mathbb{R} \to [0, \infty]$ by

$$U(\mu, x) = \limsup_{n\to\infty} \frac{n}{2}\mu((x - \frac{1}{n}, x + \frac{1}{n}]),$$

and introduce the $\mathcal{P} \times \mathcal{B}(\mathbb{R})$-measurable canonical densities on Ω^2,

$$u_i(t, X, x) = U(X_t^i, x).$$

Then

$$\begin{aligned}\Omega_{ac} &= \{X \in \Omega^2 : X_t^i \ll dx \ \forall t > 0, \ i = 1, 2\} \\ &= \{X \in \Omega^2 : X_t^i(1) = \int u_i(t, x) dx \ \forall t > 0, \ i = 1, 2\}\end{aligned}$$

is a universally measurable subset of Ω^2 (e.g. by Theorem III.4.4 (a) of Dellacherie and Meyer (1978)).

Letting $\varepsilon \downarrow 0$ in $(CS)_m^{\lambda,\varepsilon}$ suggests the following definition: A probability $\mathbb{P}$ on $(\Omega^2, \mathcal{F}^2)$ satisfies $(CS)_m^{\lambda}$ iff

$$\text{For } i = 1, 2 \ \forall \phi_i \in \mathcal{D}(\Delta/2) \quad X_t^i(\phi_i) = m_i(\phi_i) + \int_0^t X_s^i\Big(\frac{\Delta \phi_i}{2}\Big) ds$$
$$- \lambda_i \int_0^t \int \phi_i(x) u_1(s, x) u_2(s, x)\, dx\, ds + M_t^i(\phi_i),$$

where $M_t^i(\phi_i)$ are continuous $(\mathcal{F}_t^2)$ – martingales under $\mathbb{P}$ such that $M_0^i(\phi_i) = 0$,

and $\langle M^i(\phi_i), M^j(\phi_j)\rangle_t = \delta_{ij} \int_0^t X_s^i(\phi_i^2) ds.$

Recall that $M^{i,0}$ $(i = 1, 2)$ are the orthogonal martingale measures on Ω^2 under $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$–see the notation introduced at the beginning of this Chapter.

Theorem IV.2.1. Assume $d = 1$ and let

$$F = \Big\{(m_1, m_2) \in M_F(\mathbb{R})^2 : \int \int \log\Big(\frac{1}{|x_1 - x_2|}\Big)^+ dm_1(x_1) dm_2(x_2) < \infty\Big\}.$$

(a) For each $m \in F$, $(CS)_m^{\lambda}$ has a unique solution $\mathbb{P}_m^0$ given by

$$\frac{d\mathbb{P}_m^0}{d(\mathbb{P}_{m_1} \times \mathbb{P}_{m_2})}\Big|_{\mathcal{F}_t^2} = \exp\Big\{\sum_{i=1}^2 \Big[- \lambda_i \int_0^t \int u_{3-i}(s, x) dM^{i,0}(s, x)$$
$$- \frac{\lambda_i^2}{2} \int_0^t \int u_{3-i}(s, x)^2 u_i(s, x) dx ds\Big]\Big\}.$$

In particular $\mathbb{P}_m^0(\Omega_{ac}) = 1$.

(b)$(\Omega^2, \mathcal{F}^2, \mathcal{F}_t^2, X_t, (\mathbb{P}_m^0)_{m\in F})$ is a continuous Borel strong Markov process taking values in F. That is, for each $m \in F$, $\mathbb{P}_m^0(X_t \in F \ \forall t \geq 0) = 1$, $m \to \mathbb{P}_m^0$ is Borel measurable, and the $(\mathcal{F}_t^2)$-strong Markov property holds.

(c) For each $m \in F$, $\mathbb{P}_m^{\varepsilon} \overset{w}{\Rightarrow} \mathbb{P}_m^0$ as $\varepsilon \downarrow 0$.

Proof. (a) Note first that $(CS)_m^{\lambda}$ is a special case of $(MP)_g^m$ with

$$g_i(s, X, x) = \lambda_i u_{3-i}(s, X, x).$$

To see this, note that if $\mathbb{P}$ satisfies $(CS)^\lambda_m$, then by Proposition IV.1.4 (with $D^i = \theta_i = 0$) we can define a process X with law $\mathbb{P}$ and a pair of independent super-Brownian motions $(\gamma = 1)$, (Z^1, Z^2) on the same space so that $Z^i \geq X^i$. As $Z^i \ll dx$ by Theorem III.3.8(c), the same is true of X^i, and so in $(CS)^\lambda_m$, $u_i(s,x)dx = X^i_s(dx)$, and $\mathbb{P}$ satisfies $(MP)^m_g$ as claimed. The converse implication is proved in the same way. The fact that g_i can now take on the value ∞ will not alter any of the results (or proofs) in the previous section.

Now check the hypotheses of Theorem IV.1.6(c) for the above choice of g_i. Condition (IV.1.3) is obvious (with $\theta \equiv 0$). For (IV.1.1), by symmetry it suffices to show that

$$(IV.2.1) \qquad \mathbb{P}_{m_1} \times \mathbb{P}_{m_2}\Big(\int_0^t \int u_1(s,x)^2 u_2(s,x)\, dx\, ds\Big) < \infty \ \forall t > 0.$$

Recall from (III.4.1) that if $mP_t(x) = \int p_t(y-x)dm(y)$ then under $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$,

$$u_i(t,x) = m_i P_t(x) + \int_0^t \int p_{t-s}(y-x) dM^{i,0}(s,y) \quad a.s. \ \text{ for each } t, x,$$

where the stochastic integral is square integrable. This shows that

$$(IV.2.2) \qquad \mathbb{P}_{m_i}(u_i(t,x_i)) = m_i P_t(x_i),$$

and

$$\begin{aligned}
\mathbb{P}_{m_i}(u_i(t,x_i)^2) &= m_i P_t(x_i)^2 + \int_0^t \int p_{t-s}(y-x_i)^2 m_i P_s(y)\, dy\, ds \\
&\leq m_i P_t(x_i)^2 + \int_0^t (2\pi(t-s))^{-1/2}\, ds\, m_i P_t(x_i) \\
&\leq m_i P_t(x_i)^2 + \sqrt{t}\, m_i P_t(x_i) \\
(IV.2.3) \qquad &\leq m_i P_t(x_i)^2 + m_i(1).
\end{aligned}$$

Now use these estimates to bound the lefthand side of (IV.2.1) by

$$\int_0^t \int m_1 P_s(x)^2 m_2 P_s(x)\, dx\, ds + \int_0^t \int m_1(1) m_2 P_s(x)\, dx\, ds.$$

The second term is $m_1(1)m_2(1)t$ and so is clearly finite for all $t > 0$ for any pair of finite measures m. Bound $m_1 P_s(x)^2$ by $m_1(1)s^{-1/2} m_1 P_s(x)$ and use Chapman-Kolmogorov to see that the first term is at most

$$\begin{aligned}
& m_1(1) \int_0^t \int s^{-1/2} p_{2s}(y_1 - y_2) m_1(dy_1) m_2(dy_2)\, ds \\
& \leq m_1(1) \int \left(1 + \log\left(\frac{4t}{|y_1 - y_2|^2}\right)^+\right) dm_1(y_1) dm_2(y_2) \\
& < \infty \quad \text{if } m \in F.
\end{aligned}$$

(b) Let Z be the pair of independent dominating SBM's constructed in (a). Since Z^i_t has a continuous density on compact support for all $t > 0$ a.s. (Theorem III.4.2(a)

and Corollary III.1.4), clearly $Z_t \in F$ for all $t > 0$ a.s. and hence the same is true for X $\mathbb{P}^0_m$ a.s. The Borel measurability in m is clear from the Radon-Nikodym derivative provided in (a) and the strong Markov property is then a standard consequence of uniqueness (see, e.g. the corresponding discussion for $\mathbb{P}^\varepsilon_m$ at the end of the last section).
(c) Write $\mathbb{P}_m$ for $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$. Let R^ε_t be the Radon-Nikodym derivative in (IV.1.7) and R^0_t be that in (a) above. It suffices to show $R^\varepsilon_t \to R^0_t$ in $\mathbb{P}_m$-probability because as these non-negative random variables all have mean 1, this would imply L^1 convergence. To show this convergence, by symmetry it clearly suffices to prove

$$\int_0^t \int [X^2_s P_\varepsilon(x) - u_2(s,x)]^2 u_1(s,x)\, dx\, ds \to 0 \text{ in } \mathbb{P}_m\text{-probability as } \varepsilon \downarrow 0.$$

If $\delta > 0$ is fixed, the fact that $(X^2_s,\ s \geq \delta)$ has a jointly continuous uniformly bounded density shows that

$$\int_\delta^t \int [X^2_s P_\varepsilon(x) - u_2(s,x)]^2 u_1(s,x)\, dx\, ds \to 0 \text{ in } \mathbb{P}_m\text{-probability as } \varepsilon \downarrow 0.$$

Therefore it suffices to show

$$\lim_{\delta\downarrow 0} \sup_{0<\varepsilon<1} \mathbb{P}_m\Big(\int_0^\delta \int [X^2_s P_\varepsilon(x)^2 + u_2(s,x)^2] u_1(s,x)\, dx\, ds\Big) = 0.$$

The argument in (a) easily handles the $u_2(s,x)^2$ term, so we focus on the $X^2_s P_\varepsilon(x)^2$ term. Use (IV.2.2) and (IV.2.3) to see that

$$\begin{aligned}
&\mathbb{P}_m\Big(\int_0^\delta \int X^2_s P_\varepsilon(x)^2 u_1(s,x)\, dx\, ds\Big)\\
&\quad\leq \int_0^\delta \int\int p_\varepsilon(y-x) m_2 P_s(y)^2\, dy\, m_1 P_s(x)\, dx\, ds + m_2(1) m_1(1)\delta\\
&\quad\leq \int_0^\delta m_2(1) s^{-1/2} \int m_2 P_{s+\varepsilon}(x) m_1 P_s(x)\, dx\, ds + m_1(1) m_2(1)\delta\\
&\quad\to 0 \text{ as } \delta \downarrow 0,
\end{aligned}$$

by the same argument as that at the end of the proof of (a). ∎

Remark IV.2.2. In $(CS)^\lambda_m$ we may restrict the test functions ϕ_i to $C^\infty_b(\mathbb{R})$. To see this, first recall from Examples II.2.4 that this class is a core for $\mathcal{D}(\Delta/2)$. Now suppose the conclusion of $(CS)^\lambda_m$ has been verified for $C^\infty_b(\mathbb{R})$ and for a sequence of functions $\{(\phi^n_1, \phi^n_2)\}$ in $\mathcal{D}(\Delta/2)^2$ such that $(\phi^n_i, \frac{\Delta}{2}\phi^n_i) \xrightarrow{bp} (\phi_i, \frac{\Delta}{2}\phi_i)$ as $n \to \infty$ for $i = 1, 2$. It follows from $(CS)^\lambda_m$ for $\phi_i \equiv 1$ that $E(X^i_s(1)) \leq m_i(1)$. Therefore by Dominated Convergence

$$E(\sup_{t\leq T}(M^i_t(\phi^n_i) - M^i_t(\phi_i))^2) \leq cE\left(\int_0^T X^i_s((\phi^n_i - \phi_i)^2)\, ds\right) \to 0 \text{ as } n \to \infty.$$

By Dominated Convergence it is now easy to take limits in $(CS)^\lambda_m$ to see that this conclusion persists for the limiting functions (ϕ_1, ϕ_2). This establishes the claim.

We also showed in the proof of (a) above that $(CS)^\lambda_m$ is equivalent to $(MP)^m_g$ with $g_i = \lambda_i u_{3-i} \geq 0$. Hence by Remark IV.1.7(c), $(CS)^\lambda_m$ remains unchanged if we only assumed that $M^i_\cdot(\phi_i)$ are continuous $(\mathcal{F}^2_t)$-local martingales.

One can easily reformulate $(CS)^\lambda_m$ as a stochastic pde. Assume W_1, W_2 are independent white noises on $\bar\Omega = (\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$. Recall that $C_K(\mathbb{R})$ is the space of continuous functions on $\mathbb{R}$ with compact support equipped with the sup norm.

A pair of non-negative processes $\{u_1(t,x), u_2(t,x) : t > 0, x \in \mathbb{R}\}$, is a solution of

$$(\mathrm{SPDE})^\lambda_m \qquad \frac{\partial u_i}{\partial t} = \frac{\Delta u_i}{2} - \lambda_i u_1 u_2 + \sqrt{u_i}\dot W_i, \quad u_i(0+, x)dx = m_i.$$

iff for $i = 1, 2$,

(i) $\{u_i(t,\cdot) : t > 0\}$ is continuous and $(\mathcal{F}_t)$ − adapted with values in $C_K(\mathbb{R})$.

$$(ii) \langle u_i(t), \phi\rangle \equiv \int u_i(t,x)\phi(x)dx = m_i(\phi) + \int_0^t \langle u_i(s), \frac{\phi''}{2}\rangle - \lambda_i \langle u_1(s)u_2(s), \phi\rangle \, ds$$
$$+ \int_0^t \int \phi(x)\sqrt{u_i(s,x)}dW_i(s,x), \ \forall t > 0 \text{ a.s. } \forall \phi \in C^2_b(\mathbb{R}).$$

As in Remark III.4.1 this implies

$$(IV.2.4) \qquad X^i_t(dx) \equiv u_i(t,x)dx \xrightarrow{\text{a.s.}} m_i \text{ as } t \downarrow 0.$$

Proposition IV.2.3. Assume $m \in F$.
(a) If (u_1, u_2) satisfies $(\mathrm{SPDE})^\lambda_m$, and X is given by (IV.2.4), then $\mathcal{L}(X) = \mathbb{P}^0_m$. In particular, the law of u on $C((0,\infty), C_K(\mathbb{R})^2)$ is unique.
(b) There is an $\bar\Omega' = (\Omega', \mathcal{F}', \mathcal{F}'_t, \mathbb{P}')$ such that if

$$\bar\Omega = (\Omega^2 \times \Omega', \mathcal{F}^2 \times \mathcal{F}', \left(\mathcal{F}^2_\cdot \times \mathcal{F}'_\cdot\right)_{t+}, \mathbb{P}^0_m \times \mathbb{P}')$$

and $\Pi : \Omega^2 \times \Omega' \to \Omega'$ is the projection map, then there is a pair of independent white noises, $\dot W_1, \dot W_2$ on $\bar\Omega$ such that $(u_1, u_2) \circ \Pi$ solves $(\mathrm{SPDE})^\lambda_m$ on $\bar\Omega$.

Proof. (a) The weak continuity of X follows from (IV.2.4) as in the proof of Theorem III.4.2(c). It now follows from Remark IV.2.2 that X satisfies $(CS)^\lambda_m$ and hence has law $\mathbb{P}^0_m$ by Theorem IV.2.1. The second assertion now follows as in the univariate case (Corollary III.4.3(c)).

(b) Let $u_{n,i}(t, X, x) = \frac{n}{2} X^i_t((x - \frac{1}{n}, x + \frac{1}{n}])$. We know $\mathbb{P}^0_m \ll \mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$ and under the latter measure $u_i(t,x)$ is the jointly continuous density of X^i_t on $(0,\infty) \times \mathbb{R}$ (Theorem III.4.2(a)), and $\{(t,x) : u_i(t,x) > 0, t \geq \delta\}$ is bounded for every $\delta > 0$ (Corollary III.1.7). It follows that $\mathbb{P}_{m_1} \times \mathbb{P}_{m_2}$-a.s. and therefore $\mathbb{P}^0_m$-a.s. for every $\delta > 0$, and $i = 1, 2$,

$$\sup_{x\in\mathbb{R}} \sup_{t\in[\delta,\delta^{-1}]} |u_{n,i}(t,x) - u_{n',i}(t,x)|$$
$$= \sup_{x\in\mathbb{Q}} \sup_{t\in\mathbb{Q}\cap[\delta,\delta^{-1}]} |u_{n,i}(t,x) - u_{n',i}(t,x)| \to 0 \text{ as } n, n' \to \infty,$$

and $\exists R$ such that $\sup_{t\in[\delta,\delta^{-1}]} X_t^i(B(0,R)^c) = 0$. It follows that (i) holds $\mathbb{P}_m^0$-a.s. It remains to show that (ii) holds on this larger space. Choose Ω' carrying two independent white noises, W_1', W_2' on $\mathbb{R}_+ \times \mathbb{R}$. Define W_i on $\bar{\Omega}$ by

$$W_i(\omega', X)_t(A) = \int_0^t \int 1_A(x) \frac{1(u_i(s,X,x) > 0)}{\sqrt{u_i(s,X,x)}} dM^i(X)(s,x) + \int_0^t \int 1_A(x) 1(u_i(s,X,x) = 0) dW_i'(\omega')(s,x).$$

As in Theorem III.4.2(b), (W_1, W_2) are independent white noises on $\bar{\Omega}$ and $(u^1, u^2) \circ \Pi$ satisfies $(\mathrm{SPDE})_m^\lambda$ on $\bar{\Omega}$. Note the independence follows from the orthogonality of the martingales $W_1(t)(A)$ and $W_2(t)(B)$ for each A and B because these are Gaussian processes in (t, A). ∎

Here is a univariate version of the above result which may be proved in the same manner. If $\sigma^2, \gamma > 0$, $\lambda \geq 0$, and $\theta \in \mathbb{R}$, consider

$$(SPDE) \qquad \frac{\partial u}{\partial t} = \frac{\sigma^2 \Delta u}{2} + \sqrt{\gamma u}\dot{W} + \theta u - \lambda u^2, \quad u_{0+}(x)ds = m(dx),$$

where $m \in M_F(\mathbb{R})$, and the above equation is interpreted as before.

In the next result we also use $u_t(x) = \frac{dX}{dx}(x)$, to denote the canonical density of the absolutely continuous part of X_t on the canonical space of paths Ω_X (defined as before).

Proposition IV.2.4. Assume

$$\int\int \Big(\log \frac{1}{|x_1 - x_2|}\Big)^+ dm(x_1)dm(x_2) < \infty.$$

(a) There is a filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ carrying a solution of (SPDE).
(b) If u is any solution of (SPDE) and $\mathbb{P}$ is the law of $t \to u_t(x)dx$ on Ω_X, then

$$\frac{d\mathbb{P}}{d\mathbb{P}_m}\Big|_{\mathcal{F}_t^X} = \exp\Big\{\int_0^t \int (\theta - \lambda u(s,x))dM(s,x) - \frac{1}{2}\int_0^t \int (\theta - \lambda u(s,x))^2 X_s(dx)dx\Big\}.$$

Here $\mathbb{P}_m$ is the law of super-Brownian motion starting at m with spatial variance σ^2, 0 drift and branching rate γ, and $dM(s,x)$ is the associated martingale measure. In particular the law of u on $C((0,\infty), C_K(\mathbb{R})^2)$ is unique.

The above result was pointed out by Don Dawson in response to a query of Rick Durrett. Durrett's question was prompted by his conjecture that the above SPDE arises in the scaling limit of a contact process in one dimension. The conjecture was confirmed by Mueller and Tribe (1994).

3. Collision Local Time

To study $(CS)_m^\lambda$ in higher dimensions we require an analogue of $u_1(s,x)u_2(s,x)dsdx$ which will exist in higher dimensions when the measures in question will not have densities. This is the collision local time of a pair of measure-valued processes which we now define.

Definition. Let $X = (X^1, X^2)$ be a pair of continuous $M_F(\mathbb{R}^d)$-valued processes on a common probability space and let p_t denote the standard Brownian transition density. The collision local time (COLT) of X is a continuous non-decreasing $M_F(\mathbb{R}^d)$-valued process $L_t(X)$ such that for any $\phi \in C_b(\mathbb{R}^d)$ and $t \geq 0$,

$$L_t^\varepsilon(X)(\phi) \equiv \int_0^t \int \phi\left(\frac{x_1 + x_2}{2}\right) p_\varepsilon(x_1 - x_2) X_s^1(dx_1) X_s^2(dx_2) ds \xrightarrow{P} L_t(X)(\phi)$$

as $\varepsilon \to 0$.

Definition. The graph of an $M_F(\mathbb{R}^d)$-valued process $(X_t, t \geq 0)$ is

$$G(X) = \cup_{\delta > 0} \mathrm{cl}\{(t, x) : t \geq \delta,\ x \in S(X_t)\} \equiv \cup_{\delta > 0} G_\delta(X) \subset \mathbb{R}_+ \times \mathbb{R}^d.$$

Remarks IV.3.1. (a) Clearly the process $L(X)$ is uniquely defined up to null sets. It is easy to check that $L(X)(ds, dx)$ is supported by $G(X^1) \cap G(X^2)$. This random measure gauges the intensity of the space-time collisions between the populations X^1 and X^2 and so can be used as a means of introducing local interactions between these populations. See the next section and Dawson et al (2000a) for examples.

(b) If $X_s^i(dx) = u_i(s, x) dx$, where u_i is a.s. bounded on $[0, t] \times \mathbb{R}^d$, then an easy application of Dominated Convergence, shows that

$$L_t(X)(dx) = \Big(\int_0^t u_1(s, x) u_2(s, x) ds\Big) dx.$$

However $L_t(X)$ may exist even for singular measures as we will see in Theorem IV.3.2 below.

(c) The definition of collision local time remains unchanged if $L_t^\varepsilon(X)(\phi_i)$ is replaced with $L_t^{\varepsilon,i}(X)(\phi_i) = \int_0^t \int p_\varepsilon(x_1 - x_2) X_s^j(dx_j) \phi_i(x_i) X_s^i(dx_i) ds$ $(i \neq j)$. This is easy to see by the uniform continuity of ϕ on compact sets.

Throughout this Section we will assume

(H_1) $\quad Z^i$ is an $(\mathcal{F}_t) - (SBM)(\gamma_i)$ starting at $m_i \in M_F(\mathbb{R}^d),\ i = 1, 2,$ defined on $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}),$ and (Z^1, Z^2) are independent.

Let $Z_t = Z_t^1 \times Z_t^2$. Recall from Section III.5 that $g_\beta(r) = \begin{cases} r^{-\beta} & \text{if } \beta > 0 \\ 1 + \log^+ \frac{1}{r} & \text{if } \beta = 0 \\ 1 & \text{if } \beta < 0 \end{cases}.$

Theorem IV.3.2. (a) If $d \leq 5$, $m_i \neq 0$, and

$$(IC) \qquad \int g_{d-2}(|z_1 - z_2|) m_1(dz_1) m_2(dz_2) < \infty \quad \text{if } d \leq 4$$

$$\int g_{d-1}(|z_1 - z_2|) m_1(dz_1) m_2(dz_2) < \infty \quad \text{if } d = 5,$$

then $L_t(Z)$ exists, is not identically 0 and satisfies

$$\lim_{\varepsilon \downarrow 0} \| \sup_{t \leq T} |L_t^\varepsilon(Z)(\phi) - L_t(Z)(\phi)| \|_2 = 0 \quad \forall T > 0,\ \phi \in b\mathcal{B}(\mathbb{R}^d).$$

In particular, $\mathbb{P}(G(Z^1)\cap G(Z^2)\neq\emptyset)>0$.
(b) If $d\geq 6$, then $G(Z^1)\cap G(Z^2)=\emptyset$ a.s.

We will prove this below except for the critical 6-dimensional case whose proof will only be sketched.

Lemma IV.3.3. If $d\geq 2$, there is a constant $C=C(d,\gamma_1,\gamma_2)$, and for each $\delta>0$, a random $r_1(\delta,\omega)>0$ a.s. such that for all $0\leq r\leq r_1(\delta)$,

$$\sup_{t\geq\delta}\iint 1(|z_1-z_2|\leq r)Z_t^1(dz_1)Z_t^2(dz_2)\leq C(\sup_t Z_t^1(1)+1)r^{4-4/d}\Big(\log\frac{1}{r}\Big)^{2+2/d}.$$

Proof. We defer this to the end of this Section. It is a nice exercise using the results of Chapter III but the methods are not central to this Section. Clearly if $d=1$ the above supremuim is a random multiple of r by Theorem III.4.2.

Corollary IV.3.4. If $d\geq 2$ and $0\leq\beta<4-4/d$, then with probability 1,

$$\lim_{\varepsilon\downarrow 0}\sup_{t\geq\delta}\int g_\beta(|z_1-z_2|)1(|z_1-z_2|\leq\varepsilon)Z_t^1(dz_1)Z_t^2(dz_2)=0\quad\forall\delta>0$$

and

$$t\mapsto\iint g_\beta(|z_1-z_2|)Z_t^1(dz_1)Z_t^2(dz_2)\text{ is continuous on }(0,\infty).$$

In particular, this is the case for $\beta=d-2$ and $d\leq 5$.
Proof. Define a random measure on $[0,\infty)$ by

$$D_t(A)=Z_t(\{(z_1,z_2):|z_1-z_2|\in A\}).$$

If $0<\beta<4-4/d$ and $\varepsilon<r(\delta,\omega)$, then an integration by parts and Lemma IV.3.3 give

$$\begin{aligned}&\sup_{t\geq\delta}\iint g_\beta(|z_1-z_2|)1(|z_1-z_2|\leq\varepsilon)Z_t^1(dz_1)Z_t^2(dz_2)\\
&\quad=\sup_{t\geq\delta}\Big[g_\beta(r)D_t([0,r])|_{0+}^{\varepsilon}+\beta\int_0^\varepsilon r^{-1-\beta}D_t([0,r])\,dr\Big]\\
&\quad\leq C(\sup_t Z_t^1(1)+1)\Big[\varepsilon^{-\beta+4-4/d}(\log^+\frac{1}{\varepsilon})^{2+2/d}+\beta\int_0^\varepsilon r^{3-\beta-4/d}(\log^+\frac{1}{r})^{2+2/d}\,dr\Big]\\
&\quad\to 0\text{ as }\varepsilon\downarrow 0,\end{aligned}$$

by our choice of β. It follows that for all $0\leq\beta<4-4/d$,

$$\lim_{M\to\infty}\sup_{t\geq\delta}\iint(g_\beta(|z_1-z_2|)-g_\beta(|z_1-z_2|)\wedge M)\,dZ_t^1dZ_t^2=0\ \text{ a.s.}$$

(if $\beta=0$ we simply compare with a $\beta>0$). The weak continuity of Z_t shows that $t\mapsto\int g_\beta(|z_1-z_2|)\wedge M\,dZ_t$ is a.s. continuous and the second result follows. ∎

Throughout the rest of this Section we assume

(H_2) $X=(X^1,X^2)$ satisfies $(MP)^m_{C,0}$ for some C with $E_i=\mathbb{R}^d$ and $A_i=\Delta/2$ on $\bar\Omega=(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$.

Apply Proposition IV.1.4 with $D^i = 0$ to see that by enlarging the space we may assume there is a pair of independent SBM's (Z^1, Z^2) as in (H_1) defined on $\bar{\Omega}$ such that $X_t^i \le Z_t^i$ for all $t \ge 0$ and $i = 1, 2$. Set $X_t = X_t^1 \times X_t^2$. We first derive a martingale problem for X and then construct $L_t(X)$ by means of a Tanaka formula.

Notation. $\frac{\vec{\Delta}}{2}$ is the generator of the standard $2d$-dimensional Brownian motion and $\vec{P}_t$ is its semigroup.

Lemma IV.3.5. For any $\phi \in \mathcal{D}(\vec{\Delta}/2)$,

$$\begin{aligned} X_t(\phi) =& X_0(\phi) + \int_0^t \int \int \phi(x_1, x_2)[X_s^1(dx_1) M^2(ds, dx_2) + X_s^2(dx_2) M^1(ds, dx_1)] \\ & - \int_0^t \int \int \phi(x_1, x_2)[X_s^1(dx_1) C^2(ds, dx_2) + X_s^2(dx_2) C^1(ds, dx_1)] \\ & + \int_0^t X_s\Big(\frac{\vec{\Delta}\phi}{2}\Big)\, ds. \end{aligned} \tag{IV.3.1}$$

Proof. *Step 1.* $\phi(x_1, x_2) = \phi_1(x_1)\phi_2(x_2)$, $\phi_i \in \mathcal{D}(\vec{\Delta}/2)$.
Then $X_t(\phi) = X_t^1(\phi_1) X_t^2(\phi_2)$ and the result follows from $(MP)_{C,0}^m$ by an integration by parts.
Step 2. $\phi(x) = \vec{P}_\varepsilon \psi(x)$, where $\psi \in \mathcal{D}(\vec{\Delta}/2)$.
Then there is a sequence of finite Riemann sums of the form

$$\phi_n(x_1, x_2) = \sum_{y_1^{i,n}, y_2^{i,n}} p_\varepsilon(y_1^{i,n} - x_1) p_\varepsilon(y_2^{i,n} - x_2) \psi(y_1^{i,n}, y_2^{i,n}) \Delta_n,$$

such that $\phi_n \xrightarrow{bp} \phi$ and

$$\begin{aligned} \frac{\vec{\Delta}}{2}\phi_n(x_1, x_2) &= \sum_{y_1^{i,n}, y_2^{i,n}} \frac{\vec{\Delta}}{2}(p_\varepsilon(y_1^{i,n} - \cdot) p_\varepsilon(y_2^{i,n} - \cdot))(x_1, x_2) \psi(y_1^{i,n}, y_2^{i,n}) \Delta_n \\ &\xrightarrow{bp} \frac{\vec{\Delta}}{2} \vec{P}_\varepsilon \psi(x_1, x_2). \end{aligned}$$

By Step 1, (IV.3.1) holds for each ϕ_n. Now let $n \to \infty$ and use Dominated Convergence to obtain this result for ϕ.
Step 3. $\phi \in \mathcal{D}(\vec{\Delta}/2)$.
Let $\varepsilon_n \downarrow 0$ and note that $\vec{P}_{\varepsilon_n}\phi \xrightarrow{bp} \phi$ and $\frac{\vec{\Delta}}{2}\vec{P}_{\varepsilon_n}\phi = \vec{P}_{\varepsilon_n}\Big(\frac{\vec{\Delta}}{2}\phi\Big) \xrightarrow{bp} \frac{\vec{\Delta}}{2}\phi$ as $n \to \infty$. Now use (IV.3.1) for $\vec{P}_{\varepsilon_n}\phi$ (from Step 2) and let $n \to \infty$ to derive it for ϕ. ∎

Let $\phi \in C_K(\mathbb{R}^d)$ and apply the above result to $\phi_\varepsilon \in \mathcal{D}(\vec{\Delta}/2)$, which is chosen so that

$$\int_0^t X_s\Big(\frac{\vec{\Delta}}{2}\phi_\varepsilon\Big) ds = -L_t^\varepsilon(\phi). \tag{IV.3.2}$$

This will be the case if

$$(IV.3.3)\qquad \frac{\vec{\Delta}}{2}\phi_\varepsilon(x_1,x_2)=-p_\varepsilon(x_1-x_2)\phi\Big(\frac{x_1-x_2}{2}\Big)\equiv\psi_\varepsilon(x_1,x_2).$$

Let $\vec{U}_\lambda$ denote the $2d$-dimensional Brownian resolvent for $\lambda\ge 0$ and assume $d>2$. By Exercise II.2.2, $\phi_\varepsilon(x)=\vec{U}_0\psi_\varepsilon(x)\in\mathcal{D}(\vec{\Delta}/2)$ satisfies (IV.3.3). If $B_s=(B^1_s,B^2_s)$ is a $2d$-dimensional Brownian motion, then $\frac{B^1+B^2}{\sqrt{2}}$ and $\frac{B^1-B^2}{\sqrt{2}}$ are independent d-dimensional Brownian motions and so a simple calculation yields

$$\begin{aligned}\phi_\varepsilon(x_1,x_2)&=E^{x_1,x_2}\Big(\int_0^\infty p_\varepsilon(B^1_s-B^2_s)\phi\Big(\frac{B^1_s+B^2_s}{2}\Big)\,ds\Big)\\ (IV.3.4)\qquad&=2^{1-d}\int_0^\infty p_{\varepsilon/4+u}\Big(\frac{x_1-x_2}{2}\Big)P_u\phi\Big(\frac{x_1+x_2}{2}\Big)\,du\\ &\equiv G_\varepsilon\phi(x_1,x_2).\end{aligned}$$

We may use (IV.3.2) in Lemma IV.3.5 and conclude that

$$\begin{aligned}X_t(G_\varepsilon\phi)&=X_0(G_\varepsilon\phi)\\ &\quad+\int_0^t\iint G_\varepsilon\phi(x_1,x_2)[X^1_s(dx_1)M^2(ds,dx_2)+X^2_s(dx_2)M^1(ds,dx_1)]\\ (T)_\varepsilon\qquad&\quad-\int_0^t\iint G_\varepsilon\phi(x_1,x_2)[X^1_s(dx_1)C^2(ds,dx_2)+X^2_s(dx_2)C^1(ds,dx_1)]\\ &\quad-L^\varepsilon_t(\phi)\qquad\forall t>0,\quad\text{for } d>2.\end{aligned}$$

(IV.3.4) shows that $G_\varepsilon\phi$ is defined for any $\phi\in b\mathcal{B}(\mathbb{R}^d)$ and that $\phi_n\xrightarrow{bp}\phi$ implies $G_\varepsilon\phi_n\xrightarrow{bp}G_\varepsilon\phi$. Now use Dominated Convergence to extend $(T)_\varepsilon$ to all $\phi\in b\mathcal{B}(\mathbb{R}^d)$. A similar argument with

$$G_{\lambda,\varepsilon}\phi(x_1,x_2)\equiv\vec{U}_\lambda\psi_\varepsilon(x_1,x_2)=2^{1-d}\int_0^\infty e^{-2\lambda u}p_{\varepsilon/4+u}\Big(\frac{x_1-x_2}{2}\Big)P_u\phi\Big(\frac{x_1+x_2}{2}\Big)\,du$$

in place of $G_\varepsilon\phi=G_{0,\varepsilon}\phi$ shows that for any $\phi\in b\mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned}X_t(G_{\lambda,\varepsilon}\phi)&=X_0(G_{\lambda,\varepsilon}\phi)\\ &\quad+\int_0^t\iint G_{\lambda,\varepsilon}\phi(x_1,x_2)[X^1_s(dx_1)M^2(ds,dx_2)+X^2_s(dx_2)M^1(ds,dx_1)]\\ (T)_{\lambda,\varepsilon}\qquad&\quad-\int_0^t\iint G_{\lambda,\varepsilon}\phi(x_1,x_2)[X^1_s(dx_1)C^2(ds,dx_2)+X^2_s(dx_2)C^1(ds,dx_1)]\\ &\quad+\lambda\int_0^t X_s(G_{\lambda,\varepsilon}\phi)\,ds-L^\varepsilon_t(\phi)\qquad\forall t>0,\quad\text{for } d\ge 1.\end{aligned}$$

As we want to let $\varepsilon\downarrow 0$ in the above formulae, introduce

$$G_{\lambda,0}\phi(x_1,x_2)=2^{1-d}\int_0^\infty e^{-2\lambda u}p_u\Big(\frac{x_1-x_2}{2}\Big)P_u\phi\Big(\frac{x_1+x_2}{2}\Big)\,du,\quad G_0\phi=G_{0,0}\phi,$$

when this integral is well-defined, as is the case if $\phi \geq 0$. A simple integration shows that for any $\varepsilon \geq 0$,

$$(IV.3.5) \quad G_\varepsilon|\phi|(x_1,x_2) \leq \|\phi\|_\infty G_0 1(x_1,x_2) = \|\phi\|_\infty k_d g_{d-2}(|x_1-x_2|) \quad \text{if } d>2,$$

where $k_d = \Gamma(d/2-1)2^{-1-d/2}\pi^{-d/2}$. Therefore $G_0\phi(x_1,x_2)$ is finite when ϕ is bounded, $x_1 \neq x_2$, and $d>2$.

Lemma IV.3.6. Let $\phi \in b\mathcal{B}(\mathbb{R}^d)$ and $d>2$. Then

$$|G_\varepsilon\phi(x_1,x_2) - G_0\phi(x_1,x_2)| \leq \|\phi\|_\infty c_d \min(|x_1-x_2|^{2-d}, \varepsilon|x_1-x_2|^{-d}).$$

If $\phi \geq 0$, $\lim_{\varepsilon\downarrow 0} G_\varepsilon\phi(x) = G_0\phi(x)$ $(\leq \infty)$ for all x and $G_0\phi$ is lower semicontinuous.
Proof.

$$\begin{aligned}
&|G_\varepsilon\phi(x_1,x_2) - G_0\phi(x_1,x_2)| \\
&\quad \leq \|\phi\|_\infty \int_0^\infty |p_{\varepsilon/4+u}((x_1-x_2)/2) - p_u((x_1-x_2)/2)|\,du \\
&\quad \leq \|\phi\|_\infty \int_0^\infty \int_u^{u+\varepsilon/4} \Big|\frac{\partial p_v}{\partial v}((x_1-x_2)/2)\Big|\,dv\,du \\
&\quad \leq \|\phi\|_\infty \Big[\int_{\varepsilon/4}^\infty \frac{\varepsilon}{4} p_v((x_1-x_2)/2)[(x_1-x_2)^2v^{-2}/8 + d(2v)^{-1}]\,dv \\
&\qquad\qquad + \int_0^{\varepsilon/4} p_v((x_1-x_2)/2)[(x_1-x_2)^2(8v)^{-1} + d/2]dv\Big] \\
&\quad \leq \|\phi\|_\infty c'_d\varepsilon\Big[\int_0^{(x_1-x_2)^2/2\varepsilon} e^{-y}y^{d/2-1}(y+d)\,dy|x_1-x_2|^{-d} \\
&\qquad\qquad + \int_{(x_1-x_2)^2/2\varepsilon}^\infty e^{-y}y^{d/2-2}[y+(d/2)]dy|x_1-x_2|^{2-d}\Big],
\end{aligned}$$

where we substituted $y=(x_1-x_2)^2(8v)^{-1}$ in the last line . The integrand in the first term of the last line is both bounded and integrable and so the first term is at most

$$c''_d\varepsilon\|\phi\|_\infty|x_1-x_2|^{-d}\min((x_1-x_2)^2(2\varepsilon)^{-1},1).$$

The integrand in the second term is at most $c(y^{-2}\wedge 1)$ and so the second term is bounded by

$$c''_d \min(|x_1-x_2|^{2-d}, \varepsilon|x_1-x_2|^{-d}).$$

This gives the first inequality and so for the second result we need only consider $x=(x_1,x_1)$. This is now a simple consequence of Monotone Convergence. The lower semicontinuity of $G_0\phi$ follows from the fact that it is the increasing pointwise limit of the sequence of continuous functions

$$\int_{2^{-n}}^\infty p_u((x_1-x_2)/2)P_u\phi((x_1+x_2)/2)\,du. \quad \blacksquare$$

Lemma IV.3.7. If $3 \le d \le 5$, then for each $t > 0$ there is a $c_d(t)$ so that

$$\mathbb{E}\Big(\int_0^t \int \Big[\int g_{d-2}(|z_1 - z_2|) Z_s^1(dz_1)\Big]^2 Z_s^2(dz_2)ds\Big)$$
$$\le c_d(t) \iint [g_{2(d-3)}(|z_1 - z_2|) + 1]\, dm_1(z_1) dm_2(z_2).$$

Proof. We may assume $t \ge 1$. Recall that $m_2 P_s(x) = \int p_s(y-x) m_2(dy)$. Use the first and second moment calculations in Exercise II.5.2 to see that the above expectation is

$$\int_0^t \int \Big[\int g_{d-2}(|z_1 - z_2|) m_1 P_s(z_1) dz_1\Big]^2 m_2 P_s(z_2)\, dz_2 ds$$
$$+ \int_0^t \int \Big[\int_0^s \int P_{s-u}(g_{d-2}(|\cdot - z_2|))(z_1)^2 m_1 P_u(z_1) dz_1 du\Big] m_2 P_s(z_2)\, dz_2 ds$$
$$(IV.3.6) \quad \equiv I_1 + I_2.$$

Use

$$(IV.3.7) \qquad \int_0^\infty p_u(x)\, du = k(d) g_{d-2}(|x|)$$

and Chapman-Kolmogorov to see that

$$I_1 = c_d \int_0^t \int \Big[\int_0^\infty \int_u^\infty m_1 P_{s+u}(z_2) m_1 P_{s+u'}(z_2) du' du\Big] m_2 P_s(z_2) dz_2 ds$$
$$\le c_d \int_0^t \int_0^\infty \Big[\int_u^\infty (s+u')^{-d/2} du'\Big] m_1(1) \iint p_{2s+u}(z_1 - z_2) dm_1(z_1) dm_2(z_2) du ds$$
$$\le c_d m_1(1) \iint \Big[\int_0^t \int_{2s}^\infty (v-s)^{1-d/2} p_v(z_1 - z_2) dv ds\Big] dm_1(z_1) dm_2(z_2)$$
$$\le c_d m_1(1) \iint \Big[\int_0^\infty v^{1-d/2}(v \wedge t) p_v(z_1 - z_2) dv\Big] dm_1(z_1) dm_2(z_2)$$
$$\le c_d m_1(1) \iint \Big[\int_0^t v^{2-d/2} p_v(z_1 - z_2) dv + t\int_t^\infty v^{1-d} dv\Big] dm_1(z_1) dm_2(z_2).$$

A routine calculation now shows that (recall $t \ge 1$ to handle the second term)

$$(IV.3.8) \qquad I_1 \le c_d m_1(1) \begin{cases} \int (|z_1 - z_2|^{6-2d} + 1)\, dm_1(z_1) dm_2(z_2) & \text{if } d > 3 \\ \int (\log^+\Big(\frac{2t}{|z_1 - z_2|}\Big) + 1)\, dm_1(z_1) dm_2(z_2) & \text{if } d = 3. \end{cases}$$

For I_2, note first that (IV.3.7) implies

$$P_{s-u}(g_{d-2}(|\cdot - z_2|))(z_1) = k(d) \int_{s-u}^\infty p_v(z_1 - z_2)\, dv,$$

and so

$$I_2 = 2k(d)^2 \int_0^t ds \int_0^s du \int_{s-u}^\infty dv \int_v^\infty dv' \Big[\iint p_v(z_1 - z_2)p_{v'}(z_1 - z_2)$$
$$m_1 P_u(z_1) m_2 P_s(z_2) dz_1 dz_2\Big]$$
$$\le c_d \int_0^t ds \int_0^s du \int_{s-u}^\infty dv\, v^{1-d/2}\Big[\iint p_{u+v+s}(z_1 - z_2) m_1(dz_1) m_2(dz_2)\Big].$$

Use the fact that $p_{u+v+s}(x) \le 2^{d/2} p_{2(u+v)}(x)$ for $s \le u + v$ and integrate out $s \in [u,(u+v)\wedge t]$ in the above to get

$$I_2 \le c_d \iint \Big[\int_0^t \int_0^\infty v^{1-d/2} \min(v, t-u) p_{2(u+v)}(z_1 - z_2)\, dv du\Big] m_1(dz_1) m_2(dz_2)$$
$$\le c_d \iint \Big[\int_0^\infty \int_0^{w\wedge t} (w-u)^{1-d/2}((w\wedge t) - u)\, du\, p_{2w}(z_1 - z_2)\, dw\Big] m_1(dz_1) m_2(dz_2)$$
$$(IV.3.9) \quad \le c_d \iint \Big[\int_0^\infty (w\wedge t)^{3-d/2} p_{2w}(z_1 - z_2)\, dw\Big] dm_1(z_1) dm_2(z_2).$$

A change of variables now gives (recall $t \ge 1$)

$$\int_0^\infty (w\wedge t)^{3-d/2} p_{2w}(\Delta)\, dw \le c_d \Big[\Delta^{8-2d} \int_{\Delta^2/4t}^\infty x^{d-5} e^{-x} dx + t^{3-d/2} \int_t^\infty w^{-d/2} dw\Big]$$
$$\le c_d \begin{cases} \Delta^{-2} + 1 & \text{if } d = 5 \\ \log^+\left(\frac{4t}{\Delta^2}\right) + 1 & \text{if } d = 4 \\ t & \text{if } d = 3. \end{cases}$$

Use this in (IV.3.9) to see that

$$I_2 \le c_d(t) \begin{cases} \iint (|z_1 - z_2|^{-2} + 1)\, m_1(dz_1) m_2(dz_2) & \text{if } d = 5 \\ \iint (\log^+\left(\frac{1}{|z_1 - z_2|}\right) + 1)\, m_1(dz_1) m_2(dz_2) & \text{if } d = 4 \\ m_1(1) m_2(1) & \text{if } d = 3. \end{cases}$$

Combine this with (IV.3.8) and (IV.3.6) to complete the proof. ∎

Theorem IV.3.8. Assume X satisfies (H_2) where $d \le 5$ and m_1, m_2 satisfy (IC).
(a) $L_t(X)$ exists and for any $\phi \in b\mathbb{B}(\mathbb{R}^d)$,

$$(IV.3.10) \qquad \sup_{t\le T} |L_t^\varepsilon(X)(\phi) - L_t(X)(\phi)| \xrightarrow{L^2} 0 \text{ as } \varepsilon \downarrow 0 \quad \text{for all } T > 0.$$

(b) If $\lambda = 0$ and $d \ge 3$, or $\lambda > 0$ and $d \ge 1$, then for any $\phi \in b\mathcal{B}(\mathbb{R}^d)$,

$$
\begin{aligned}
X_t(G_{\lambda,0}\phi) = X_0(G_{\lambda,0}\phi) + \int_0^t \iint G_{\lambda,0}\phi(x_1,x_2)[X_s^1(dx_1)M^2(ds,dx_2) \\
+ X_s^2(dx_2)M^1(ds,dx_1)] \\
(T) \qquad - \int_0^t \iint G_{\lambda,0}\phi(x_1,x_2)[X_s^1(dx_1)C^2(ds,dx_2) \\
+ X_s^2(dx_2)C^1(ds,dx_1)] \\
+ \lambda \int_0^t X_s(G_{\lambda,0}\phi)\,ds - L_t(X)(\phi) \quad \forall t \geq 0 \quad \text{a.s.}
\end{aligned}
$$

Each of the above processes are a.s. continuous in $t \geq 0$. The second term on the right-hand side is an L^2 $(\mathcal{F}_t)$-martingale and each of the other processes on the right-hand side has square integrable total variation on bounded time intervals.

Proof. We give the proof for $\lambda = 0$ and $d \geq 3$. The extra term involved when $\lambda > 0$ is very easy to handle and for $d \leq 3$ the entire proof simplifies considerably by means of a systematic use of Theorem III.3.4 (the reader may want to consider only this case, especially as the interactive models discussed in the next Section will only exist in these dimensions).

Let $\phi \in b\mathcal{B}(\mathbb{R}^d)_+$ and note that it suffices to prove the Theorem for such a non-negative ϕ. Consider the limit of each of the terms in $(T)_\varepsilon$ as $\varepsilon \downarrow 0$. (IC) and (IV.3.5) allow us to use Dominated Convergence and conclude from Lemma IV.3.6 that

$$
(IV.3.11) \qquad \lim_{\varepsilon \downarrow 0} \iint G_\varepsilon \phi dm_1 dm_2 = \iint G_0 \phi dm_1 dm_2.
$$

Let

$$
N_t^\varepsilon(\phi) = \int_0^t \iint G_\varepsilon \phi(x_1,x_2)[X_s^1(dx_1)M^2(ds,dx_2) + X_s^2(dx_2)M^1(ds,dx_1)], \quad \varepsilon, t \geq 0.
$$

Note that Lemma IV.3.7, (IV.3.5) and the domination $X^i \leq Z^i$ show that $N_t^\varepsilon(\phi)$ is a well-defined continuous square-integrable martingale even for $\varepsilon = 0$. Similarly, Lemmas IV.3.6 and IV.3.7, this domination, and Dominated Convergence show that for any $T > 0$,

$$
\begin{aligned}
&\mathbb{E}(\sup_{t \leq T}(N_t^\varepsilon(\phi) - N_t^0(\phi))^2) \\
(IV.3.12) \qquad &\leq c\mathbb{E}\Big(\int_0^T \gamma_2 \Big(\int |G_\varepsilon\phi - G_0\phi|(x_1,x_2) Z_s^1(dx_1)\Big)^2 Z_s^2(dx_2) ds \\
&\quad + \int_0^T \gamma_1 \Big(\int |G_\varepsilon\phi - G_0\phi|(x_1,x_2) Z_s^2(dx_2)\Big)^2 Z_s^1(dx_1) ds\Big) \\
&\to 0 \text{ as } \varepsilon \downarrow 0.
\end{aligned}
$$

If $C(ds,dx_1,dx_2) = X_s^1(dx_1)C^2(ds,dx_2) + X_s^2(dx_2)C^1(ds,dx_1)$ then $(T)_\varepsilon$ implies for any $t>0$,

$$\int_0^t \int G_\varepsilon 1(x_1,x_2)C(ds,dx_1,dx_2) \le m_1 \times m_2(G_\varepsilon 1) + N_t^\varepsilon(1)$$
$$\xrightarrow{L^2} m_1 \times m_2(G_0 1) + N_t^0(1),$$

the last by (IV.3.11) and (IV.3.12). Fatou's lemma and the equality in (IV.3.5) now show that

$$(IV.3.13) \qquad \mathbb{E}\Big(\Big(\int_0^t \iint g_{d-2}(|x_1-x_2|)C(ds,dx_1,dx_2)\Big)^2\Big) < \infty \quad \forall t>0.$$

This allows us to apply Lemma IV.3.6 and Dominated Convergence to conclude

$$(IV.3.14) \qquad \lim_{\varepsilon\downarrow 0}\mathbb{E}\Big(\int_0^t \iint |G_\varepsilon\phi - G_0\phi(x_1,x_2)|C(ds,dx_1,dx_2)^2\Big) = 0.$$

$(T)_\varepsilon$ shows that $X_t(G_\varepsilon\phi) \le X_0(G_\varepsilon\phi) + N_t^\varepsilon(\phi)$ for all $t\ge 0$ a.s. Let $\varepsilon \downarrow 0$, use Lemma IV.3.6 and Fatou's Lemma on the left-hand side, and (IV.3.11) and (IV.3.12) on the right-hand side to see that

$$(IV.3.15) \qquad X_t(G_0\phi) \le X_0(G_0\phi) + N_t^0(\phi) < \infty \quad \forall t \ge 0 \text{ a.s.}$$

Take $\phi = 1$ in the above inequality, recall that $N_t^0(1)$ is an L^2-martingale, and use the equality in (IV.3.5) to get

$$(IV.3.16) \qquad \mathbb{E}\Big(\Big(\sup_{t\le T}\iint g_{d-2}(|x_1-x_2|)X_t^1(dx_1)X_t^2(dx_2)\Big)^2\Big) < \infty \quad \forall T>0.$$

The bound in Lemma IV.3.6 shows that for any $T,\ \delta,\ \eta>0$, if

$$S_\delta = \{(x_1,x_2)\in\mathbb{R}^{2d} : |x_1-x_2|\le\delta\},$$

then

$$\begin{aligned}
&\sup_{t\le T} X_t(|G_\varepsilon\phi - G_0\phi|)\\
&\quad\le \sup_{t\le T} X_t(|G_\varepsilon\phi - G_0\phi|1_{S_\delta^c})\\
(IV.3.17) \qquad &\quad + c_d\|\phi\|_\infty \sup_{t\le T}\int g_{d-2}(|x_1-x_2|)1_{S_\delta}(x_1,x_2)X_t(dx_1,dx_2)\\
&\quad\le c_d\|\phi\|\Big[\varepsilon\delta^{-d}\sup_{t\le T}X_t(1) + \sup_{\eta\le t\le T}\int g_{d-2}(|x_1-x_2|)1_{S_\delta}(x_1,x_2)X_t(dx_1,dx_2)\\
&\qquad + \sup_{t<\eta}\int g_{d-2}(|x_1-x_2|)1_{S_\delta}(x_1,x_2)X_t(dx_1,dx_2)\Big].
\end{aligned}$$

Write $X_t(g_{d-2})$ for $\int g_{d-2}(|x_1-x_2|)X_t(dx_1,dx_2)$. The lower semicontinuity of $(x_1,x_2)\to g_{d-2}(|x_1-x_2|)$ (take $\phi=1$ in Lemma IV.3.6 and use the equality in (IV.3.5)) and the weak continuity of X show that $\liminf_{t\downarrow 0} X_t(g_{d-2}) \ge X_0(g_{d-2})$.

On the other hand (IV.3.15) with $\phi = 1$ implies $\limsup_{t\downarrow 0} X_t(g_{d-2}) \le X_0(g_{d-2})$ a.s., and so

$$(IV.3.18) \qquad \lim_{t\downarrow 0} X_t(g_{d-2}) = X_0(g_{d-2}) \text{ a.s.}$$

Choose $\delta_n \downarrow 0$ so that $X_0(\{(x_1,x_2) : |x_1 - x_2| = \delta_n\}) = 0$. Weak continuity then implies $\lim_{t\downarrow 0} X_t(g_{d-2}1_{S^c_{\delta_n}}) = X_0(g_{d-2}1_{S^c_{\delta_n}})$ and so (IV.3.18) gives

$$(IV.3.19) \qquad \lim_{t\downarrow 0} X_t(g_{d-2}1_{S_{\delta_n}}) = X_0(g_{d-2}1_{S_{\delta_n}}) \text{ a.s.}$$

Let $\varepsilon_0 > 0$ and first choose an natural number N_0 so that the right-hand side is at most ε_0 for $n \ge N_0$. Next use (IV.3.19) to choose $\eta = \eta(\varepsilon_0)$ so that

$$(IV.3.20) \qquad \forall n \ge N_0 \quad \sup_{t<\eta} X_t(g_{d-2}1_{S_{\delta_n}}) \le \sup_{t<\eta} X_t(g_{d-2}1_{S_{\delta_{N_0}}}) < 2\varepsilon_0.$$

By Corollary IV.3.4 we may omit a $\mathbb{P}$-null set and then choose $N_1(\eta) \ge N_0$ so that

$$(IV.3.21) \qquad \sup_{t\ge\eta} X_t(g_{d-2}1_{S_{\delta_{N_1}}}) < \varepsilon_0.$$

Now take $\delta = \delta_{N_1}$ and $\eta = \eta(\varepsilon_0)$ in (IV.3.17). By (IV.3.20) and (IV.3.21) we see that outside a null set for $\varepsilon < \varepsilon(\varepsilon_0)$, the right-hand side of (IV.3.17) will be at most $\|\phi\|_\infty c_d 4\varepsilon_0$. We have proved

$$(IV.3.22) \qquad \lim_{\varepsilon\downarrow 0} \sup_{t\le T} X_t(|G_\varepsilon\phi - G_0\phi|) = 0 \quad \forall T > 0 \text{ a.s. and in } L^2,$$

where Dominated Convergence, Lemma IV.3.6, and (IV.3.16) are used for the L^2-convergence.

(IV.3.11), (IV.3.12), (IV.3.14) and (IV.3.22) show that each term in $(T)_\varepsilon$, except perhaps for $L^\varepsilon_t(\phi)$, converges uniformly in compact time intervals in L^2. Therefore there is an a.s. continuous process $\{\tilde{L}_t(\phi) : t \ge 0\}$, so that

$$(IV.3.23) \qquad \lim_{\varepsilon\downarrow 0} \|\sup_{t\le T} |L^\varepsilon_t(\phi) - \tilde{L}_t(\phi)|\|_2 = 0 \quad \forall T > 0.$$

Take L^2 limits uniformly in $t \le T$ in $(T)_\varepsilon$ to see that
(IV.3.24)

$$X_t(G_0\phi) = m_1\times m_2(\phi) + N^0_t(\phi) - \int_0^t G_0\phi(x_1,x_2)C(ds,dx_1,dx_2) - \tilde{L}_t(\phi) \quad \forall t \ge 0 \text{ a.s.},$$

where each term is a.s. continuous in t, $N^0_t(\phi)$ is an L^2 martingale and the last two terms have square integrable total variation on compact time intervals.

To complete the proof we need to show there is a continuous increasing $M_F(\mathbb{R}^d)$-valued process $L_t(X)$ such that

$$(IV.3.25) \qquad L_t(X)(\phi) = \tilde{L}_t(\phi) \quad \forall t \ge 0 \text{ a.s. for all } \phi \in b\mathcal{B}(\mathbb{R}^d)_+.$$

Note that (IV.3.23) then identifies $L(X)$ as the collision local time of X as the notation suggests. Let D_0 be a countable dense set in

$$C_\ell(\mathbb{R}^d) = \{\phi \in C_b(\mathbb{R}^d) : \phi \text{ has a limit at } \infty\}$$

containing 1. Choose $\varepsilon_n \downarrow 0$ and ω outside a null set so that

$$(IV.3.26) \qquad \lim_{n\to\infty} \sup_{t\le n} |L_t^{\varepsilon_n}(\phi) - \tilde{L}_t(\phi)| = 0 \quad \text{for all } \phi \in D_0,$$

and (recall Corollary III.1.7)

$$(IV.3.27) \qquad \mathcal{R}_\delta \equiv \mathrm{cl}(\cup_{t\ge\delta} S(Z_t^1) \cup S(Z_t^2)) \text{ is compact for all } \delta > 0.$$

Let $\eta > 0$. The definition of L^ε shows that $K_\delta = \left\{\frac{x_1+x_2}{2} : x_i \in \mathcal{R}_\delta\right\}$ is a compact support for $L_\infty^\varepsilon(X) - L_\delta^\varepsilon(X)$. Our choice of ω implies $\tilde{L}_\cdot(1)$ is continuous and allows us to choose $\delta > 0$ so that $L_\delta^{\varepsilon_n}(1) < \eta$ for all n. Therefore

$$L_\infty^{\varepsilon_n}(X)(K_\delta^c) = L_\delta^{\varepsilon_n}(K_\delta^c) < \eta \quad \text{for all } n.$$

Therefore $\{L_t^{\varepsilon_n}(X) : n \in \mathbb{N},\ t \ge 0\}$ are tight and (IV.3.26) shows that for each $t \ge 0$, all limit points of $\{L_t^{\varepsilon_n}\}$ in the weak topology on $M_F(\mathbb{R}^d)$ coincide. Therefore there is an $M_F(\mathbb{R}^d)$-valued process $L_t(X)$ such that $\lim_{n\to\infty} L_t^{\varepsilon_n}(X) = L_t(X)$ for all $t \ge 0$ a.s., $L_t(X)$ is non-decreasing in t and satisfies

$$(IV.3.28) \qquad L_t(X)(\phi) = \tilde{L}_t(\phi) \text{ for all } t \ge 0 \text{ and } \phi \in D_0 \text{ a.s.}$$

In particular $L_t(X)(\phi)$ is continuous in $t \ge 0$ for all $\phi \in D_0$ a.s. and hence $L_t(X)$ is a.s. continuous in t as well. If $\psi_n \xrightarrow{bp} \psi$, then using Dominated Convergence in (IV.3.24) one can easily show there is a subsequence such that

$$\lim_{k\to\infty} \tilde{L}_t(\psi_{n_k}) = \tilde{L}_t(\psi)\ \forall t \ge 0 \text{ a.s.}$$

by showing this is the case for each of the other terms in (IV.3.24). (A subsequence is needed as one initially obtains L^2 convergence for the martingale terms.) It then follows from (IV.3.28) that (IV.3.25) holds and the proof is complete. ∎

Proof of Theorem IV.3.2. (a) As we may take $X = Z$ in Theorem IV.3.8, it remains only to show that $L_t(Z)$ is not identically 0. The L^2 convergence in Theorem IV.3.8 and a simple second moment calculation show that

$$\mathbb{E}(L_t(Z)(1)) = \lim_{\varepsilon\downarrow 0} \mathbb{E}(L_t^\varepsilon(Z)(1)) = \frac{1}{2}\iint \int_0^{2t} p_s(z_1 - z_2)ds\, m_1(dz_1)m_2(dz_2) \neq 0.$$

(b) We first give a careful argument if $d > 6$. Recall the definition of $G_\delta(X)$ given at the beginning of this Section and recall that $h(u) = (u\log^+(1/u))^{1/2}$. If $\delta_i(3,\omega)$ is as in Corollary III.1.5, then that result and the fixed time hitting estimate, Theorem III.5.11, show that for $x \in \mathbb{R}^d$ and $t > 0$,

$$\begin{aligned}
&\mathbb{P}(Z_s^i(B(x,\varepsilon)) > 0 \text{ for some } s \in [t, t+\varepsilon^2(\log^+(1/\varepsilon))^{-1}],\\
&\qquad\qquad\qquad\text{and } \delta_i(3,\omega) > \varepsilon^2(\log^+(1/\varepsilon))^{-1})\\
&\le \mathbb{P}(Z_t^i(B(x, \varepsilon + 3h(\varepsilon^2(\log^+(1/\varepsilon))^{-1}))) > 0)\\
&\le C_d \gamma_i^{-1} t^{-d/2} m_i(1)\Big(\varepsilon + 3h(\varepsilon^2(\log^+(1/\varepsilon))^{-1})\Big)^{d-2}\\
(IV.3.29) \qquad &\le C_d' \gamma_i^{-1} t^{-d/2} m_i(1)\varepsilon^{d-2}.
\end{aligned}$$

Let $S_n = \{B(x_i^n, 2^{-n}) : 1 \le i \le c_d n^d 2^{nd}\}$ be an open cover of $[-n,n]^d$. If $\delta > 0$, $\mathcal{R}_\delta$ is as in (IV.3.27), and $\eta_n = 2^{-2n}(\log 2^n)^{-1}$, then

$$\begin{aligned}
&\mathbb{P}(G_\delta(Z^1) \cap G_\delta(Z^2) \ne \emptyset, \mathcal{R}_\delta \subset [-n,n]^d, Z_n^1 = Z_n^2 = 0, \delta_1(3) \wedge \delta_2(3) > \eta_n,) \\
&\quad \le \sum_{0 \le j \le n\eta_n^{-1}} \sum_{1 \le i \le c_d n^d 2^{nd}} \mathbb{P}(Z_s^1(B(x_i^n, 2^{-n}))Z_s^2(B(x_i^n, 2^{-n})) > 0 \\
&\qquad\qquad \text{for some } s \in [\delta + j\eta_n, \delta + (j+1)\eta_n], \delta_1(3) \wedge \delta_2(3) > \eta_n) \\
&\quad \le (n\eta_n^{-1} + 1) c_d n^d 2^{nd} (C_d')^2 (\gamma_1\gamma_2)^{-1} \delta^{-d} m_1(1) m_2(1) 2^{-n2(d-2)} \qquad \text{by (IV.3.29)} \\
&\quad \le c(d,\delta) m_1(1) m_2(1) n^{2+d} 2^{-n(d-6)} \to 0 \text{ as } n \to \infty.
\end{aligned}$$

As $n \to \infty$ the left-hand side of the above approaches $\mathbb{P}(G_\delta(Z^1) \cap G_\delta(Z^2) \ne \emptyset)$ by Corollary III.1.5 and (IV.3.27), and so the result follows by letting $\delta \downarrow 0$.

Finally we sketch the argument in the critical 6-dimensional case. First (IV.3.29) can be strengthened to

$$(IV.3.30) \quad \mathbb{P}\Big(\int_t^{t+\varepsilon^2} Z_s^i(B(x,\varepsilon))\, ds > 0\Big) \le c_d t^{-d/2} m_i(1) \varepsilon^{d-2} \quad \forall t \ge 4\varepsilon^2,\ d \ge 3.$$

This may shown using an appropriate nonlinear pde as in Section III.5. A short proof is given in Proposition 3.3 of Barlow, Evans and Perkins (1991). Now introduce a restricted Hausdorff measure $q^f(A)$ for $A \subset \mathbb{R}_+ \times R^d$ and $f : [0,\varepsilon) \to \mathbb{R}_+$ a non-decreasing function for which $f(0+) = 0$. It is given by

$$q^f(A) = \liminf_{\delta \downarrow 0} \Big\{ \sum_{i=1}^\infty f(r_i) : A \subset \cup_{i=1}^\infty [t_i, t_i + r_i^2] \times \prod_{j=1}^d [x_i^j, x_i^j + r_i],\ r_i < \delta \Big\}.$$

If $d > 4$ and $\psi_d(r) = r^4 \log\log(1/r)$ (as in Theorem III.3.9) then there are $0 < c_1(d) \le c_2(d) < \infty$ so that

$$\begin{aligned}
&c_1 q^{\psi_d}(A \cap G(Z^i)) \le \int_0^\infty \int 1_A(s,x) Z_s^i(dx)\, ds \le c_2 q^{\psi_d}(A \cap G(Z^i)) \\
&(IV.3.31) \qquad\qquad \forall A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d) \text{ a.s.}
\end{aligned}$$

This is Theorem 3.1 of Barlow, Evans, and Perkins (1991) and may be shown using the ideas presented in Section III.3. (It should be possible to prove $c_1 = c_2$ here.) If $q^d = q^{r^d}$, then a simple consequence of (IV.3.30) (cf. Corollary III.5.10) is

$$(IV.3.32) \quad q^{d-2}(A) = 0 \text{ implies } A \cap G(Z^1) = \emptyset \text{ a.s. for all } A \subset \mathbb{R}_+ \times \mathbb{R}^d, d \ge 3.$$

(IV.3.31) shows that $q^{d-2}(G(Z^2)) = 0$ if $d \ge 6$ and so (IV.3.32) with $A = G(Z^2)$ implies that $G(Z^1) \cap G(Z^2) = \emptyset$ a.s. ∎

Proof of Lemma IV.3.3. If $d = 2$ this is a simple consequence of Theorem III.3.4, so assume $d > 2$. We may assume that our space carries independent historical processes (H^1, H^2) associated with (Z^1, Z^2). Let h and $\delta_i(3, H^i)$ be as in the Historical Modulus of Continuity (Theorem III.1.3) and let $\bar h_d$ and $r_0(\delta, H^1)$ be as in Theorem III.3.4. Those results show that on

$$\{\omega : \delta_i(3, H^i) > 2^{-n},\ i = 1,2,\ \text{and}\ r_0(\delta, H^1) > h(2^{-n})\},$$

we have

$$\begin{aligned}
&\sup_{t\ge\delta, t\in[j2^{-n},(j+1)2^{-n}]} \iint 1(|z_1 - z_2| \le h(2^{-n}))Z_t^1(dz_1)Z_t^2(dz_2) \\
&\le \sup_{t\ge\delta, t\in[j2^{-n},(j+1)2^{-n}]} \iint 1(|y_1(j2^{-n}) - y_2(j2^{-n}| \le 7h(2^{-n}), \\
&\qquad\qquad |y_1(t) - y_2(t)| \le h(2^{-n}))H_t^1(dy_1)H_t^2(dy_2) \\
&\le \sup_{t\in[j2^{-n},(j+1)2^{-n}]} \int \gamma_1 c(d)\bar h_d(h(2^{-n})) \\
&\qquad\qquad \times 1\Big(y_2(j2^{-n}) \in S(Z^1_{j2^{-n}})^{7h(2^{-n})}\Big)H_t^2(dy_2).
\end{aligned} \tag{IV.3.33}$$

A weak form of Lemma III.1.6 (with s fixed) has also been used in the last line. If $H^{1,*} = \sup_t H_t^1(1)$, then (III.3.1) and the ensuing calculation show that for $n \ge N(H^1)$,

$$\begin{aligned}
&S(Z^1_{j2^{-n}})^{7h(2^{-n})} \\
&\subset \text{a union of } \gamma_1^{-1}(H^{1,*}+1)2^{n+2} \text{ balls of radius } 10h(2^{-n})\ \forall j \in \mathbb{N}.
\end{aligned} \tag{IV.3.34}$$

Let $W_n(j) = Z^2_{j2^{-n}}(S(Z^1_{j2^{-n}})^{7h(2^{-n}})$. Condition on H^1 and assume that $n \ge N(H^1)$. Then (IV.3.34) implies

$$P^x(B_s \in S(Z^1_{j2^{-n}})^{7h(2^{-n}}) \le c_d\gamma_1^{-1}(H^{1,*}+1)2^n h(2^{-n})^d s^{-d/2}. \tag{IV.3.35}$$

Therefore

$$\begin{aligned}
&\int_0^\infty \sup_x P^x(B_s \in S(Z^1_{j2^{-n}})^{7h(2^{-n})})\,ds \\
&\qquad \le (c_d\gamma_1^{-1}+1)(H^{1,*}+1)\int_0^\infty \min(2^n h(2^{-n})^d s^{-d/2}, 1)\,ds \\
&\qquad \le c(d,\gamma_1)(H^{1,*}+1)2^{-n(1-2/d)}\log 2^n \\
&\qquad \equiv \gamma_2^{-1}\lambda_n^{-1}.
\end{aligned}$$

If $f_n(x) = \lambda_n 1(x \in S(Z^1_{j2^{-n}})^{7h(2^{-n})})$ and $G(f_n,t)$ is as in Lemma III.3.6, then $\gamma_2 G(f_n, j2^{-n}) \le 1$ and so Lemma III.3.6 implies that on $\{n \ge N(H^1)\}$ and for $j2^{-n} \ge 1/n$,

$$\begin{aligned}
\mathbb{P}(W_n(j) \ge 17n\lambda_n^{-1}|H^1) &\le e^{-17n}\mathbb{E}(e^{\lambda_n W_n(j)}|H^1) \\
&\le e^{-17n}\exp\Big(m_2(1)2\lambda_n \sup_x P^x(B_{j2^{-n}} \in S(Z^1_{j2^{-n}})^{7h(2^{-n})})\Big) \\
&\le e^{-17n}\exp(m_2(1)c'(\delta,\gamma_1,\gamma_2)) \quad (\text{by } (IV.3.35)).
\end{aligned} \tag{IV.3.36}$$

The Markov property and (III.1.3) show that conditional on $\sigma(H^1)\vee\sigma(H^2_s, s\le j2^{-n})$, $t\mapsto H^2_{j2^{-n}+t}(\{y: y(j2^{-n})\in S(Z^1_{j2^{-n}})^{7h(2^{-n})}\})$ is equal in law to $\mathbb{P}_{W_n(j)\delta_0}(Z^2_\cdot(1)\in\cdot)$. Therefore if $\eta_n>0$ and $K_n=17n\lambda_n^{-1}$, then

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[j2^{-n},(j+1)2^{-n}]} H^2_t(\{y: y(j2^{-n})\in S(Z^1_{j2^{-n}})^{7h(2^{-n})}\})>\eta_n,\ W_n(j)<K_n|H^1\Big)\\
&\le \mathbb{E}\Big(\mathbb{P}_{W_n(j)\delta_0}\Big(\sup_{t\le 2^{-n}}\exp(2^n\gamma_2^{-1}Z^2_t(1))>\exp(2^n\eta_n\gamma_2^{-1})\Big)1(W_n(j)<K_n)|H^1\Big)\\
&\le \exp(-2^n\eta_n\gamma_2^{-1})\mathbb{E}\Big(\mathbb{E}_{K_n\delta_0}(\exp(2^n\gamma_2^{-1}Z^2_{2^{-n}}(1)))|H^1\Big)\quad\text{(weak } L^1 \text{ inequality)}\\
&\le \exp(-2^n\eta_n\gamma_2^{-1})\mathbb{E}(\exp(K_n2^{n+1}/\gamma_2)|H^1)\qquad\text{(Lemma III.3.6).}
\end{aligned}$$

Set $\eta_n=35n/\lambda_n=c''(d,\gamma_1,\gamma_2)(H^{1,*}+1)2^{-n(1-2/d)}n^2$ and use (IV.3.36) in the above to conclude that on $\{n\ge N(H^1)\}$ and for $j2^{-n}\ge 1/n$,

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{t\in[j2^{-n},(j+1)2^{-n}]} H^2_t(\{y: y(j2^{-n})\in S(Z^1_{j2^{-n}})^{7h(2^{-n})}\})>\eta_n|H^1\Big)\\
&\le e^{-17n}\exp(m_2(1)c'(\delta,\gamma_1,\gamma_2))+\exp\Big(-\frac{2^n\eta_n}{\gamma_2}+\frac{(34)2^nn}{\gamma_2\lambda_n}\Big)\\
&\le e^{-17n}\exp(m_2(1)c'(\delta,\gamma_1,\gamma_2))+\exp\Big(-\frac{2^nn}{\gamma_2\lambda_n}\Big)\\
&\le e^{-17n}\exp(m_2(1)c'(\delta,\gamma_1,\gamma_2))+\exp(-c(d,\gamma_1)2^{2n/d}(\log 2)n^2).
\end{aligned}$$

A conditional application of Borel-Cantelli now shows there is an $N(H)<\infty$ a.s. so that for $n\ge N(H)$,

$$\begin{aligned}
&\sup_{j2^{-n}\ge 1/n}\ \sup_{t\in[j2^{-n},(j+1)2^{-n}]} H^2_t(\{y: y(j2^{-n})\in S(Z^1_{j2^{-n}})^{7h(2^{-n})}\})\\
&\qquad\le c''(\delta,\gamma_1,\gamma_2)(H^{1,*}+1)2^{-n(1-2/d)}n^2.
\end{aligned}$$

Use this in (IV.3.33) to see that for a.a. ω if n is sufficiently large, then

$$\begin{aligned}
&\sup_{t\ge\delta}\iint 1(|z_1-z_2|\le h(2^{-n}))Z^1_t(dz_1)Z^2_t(dz_2)\\
&\qquad\le \gamma_1c(d)c''(d,\gamma_1,\gamma_2)(H^{1,*}+1)\bar{h}_d(h(2^{-n}))2^{-n(1-2/d)}n^2.
\end{aligned}$$

An elementary calculation now completes the proof. ∎

4. A Singular Competing Species Model–Higher Dimensions.

In this Section we describe how to use collision local time to formulate and solve the competing species model introduced in Section IV.1 in higher dimensions. The actual proof of the main results (due to Evans and Perkins (1994,1998) and Mytnik (1999)) are too long to reproduce here and so this Section will be a survey of known results together with some intuitive explanations.

We use the notation of Section IV.1 with $E_i=\mathbb{R}^d$ and $A_i=\Delta/2$. In particular, $\Omega^2=C(\mathbb{R}_+,M_F(\mathbb{R}^d))^2$ with its Borel σ-field $\mathcal{F}^2$ and canonical right-continuous filtration $\mathcal{F}_t^2$. In view of Remark IV.3.1(b), here is the natural extension of $(CS)_m^\lambda$ to higher dimensions.

Definition. Let $\lambda = (\lambda_1, \lambda_2) \in \mathbb{R}_+^2$ and $m = (m_1, m_2) \in M_F(\mathbb{R}^d)^2$. A probability $\mathbb{P}$ on $(\Omega^2, \mathcal{F}^2)$ satifies $(GCS)_m^\lambda$ iff

$$X_t^i(\phi_i) = m_i(\phi_i) + \int_0^t X_s^i\Big(\frac{\Delta\phi_i}{2}\Big)ds + M_t^i(\phi_i) - \lambda_i L_t(X)(\phi_i),$$

where $M_t^i(\phi_i)$ is a continuous $\mathcal{F}_t^2$ − martingale under $\mathbb{P}$ such that

$$(GCS)_m^\lambda \quad M_0^i(\phi_i) = 0 \text{ and} \langle M^i(\phi_i), M^j(\phi_j)\rangle_t = \delta_{ij} \int_0^t X_s^i(\phi_i^2)ds$$

for $\phi_i \in \mathcal{D}(\Delta/2)$ and $i = 1, 2$.

The existence of $L_t(X)$ is implicit in $(GCS)_m^\lambda$. We will say $X = (X_1, X_2)$ satisfies $(GCS)_m^\lambda$ if X is a process whose law satisfies $(GCS)_m^\lambda$. Let $(GCS)_m^{\varepsilon,\lambda}$ denote the corresponding martingale problem in which $L_t(X)(\phi_i)$ is replaced by $L_t^{\varepsilon,i}(X)(\phi_i)$ for $i = 1, 2$ (recall Remark IV.3.1(c)).

Note first that the Domination Principle (Proposition IV.1.4) shows that if X satisfies $(CS)_m^\lambda$, we may assume there are a pair of independent super-Brownian motions (Z^1, Z^2) such that $X^i \le Z^i$ a.s. If $d \ge 6$, then Theorem IV.3.2(b) implies $G(X^1) \cap G(X^2) \subset G(Z^1) \cap G(Z^2) = \emptyset$ and so $L(X)$ must be 0 a.s. Corollary IV.1.3 (with $g_i^0 = 0$) now shows that (X^1, X^2) is a pair of independent super-Brownian motions. Conversely, using Theorem IV.3.2(b) it is easy to see that a pair of independent super-Brownian motions does satisfy $(GCS)_m^\lambda$ with $L(X) = 0$ if $\int g_{d-2}(|z_1 - z_2|)m_1(dz_1)m_2(dz_2) < \infty$. (The latter condition ensures that $\sup_{\varepsilon>0} \mathbb{E}(L_\delta^\varepsilon(X)(1))$ approaches 0 as $\delta \downarrow 0$, and Theorem IV.3.2(b) shows that $L_t^\varepsilon(X)(1) - L_\delta^\varepsilon(X)(1) \overset{a.s.}{\to} 0$ as $\varepsilon \downarrow 0$ for any $\delta > 0$.) Therefore we only consider the above martingale problem for $d \le 5$ when non-trivial solutions may exist.

Next we show that if $d = 1$, then $(GCS)_m^\lambda$ may be viewed as a generalization of $(CS)_m^\lambda$.

Proposition IV.4.1. Assume $d = 1$ and $m \in F$. The unique solution P_m^0 of $(CS)_m^\lambda$ also satisfies $(GCS)_m^\lambda$.

Proof. We need only show that

$$(IV.4.1) \qquad L_t(X)(dx) = \Big(\int_0^t u_1(s,x)u_2(s,x)ds\Big)dx \quad \mathbb{P}_m^0\text{a.s.}$$

Let $\phi \in C_b(\mathbb{R})$. Theorem IV.2.1 shows that $X_s^i(dx) = u_i(s,x)dx$ for all $s > 0$ $\mathbb{P}_m^0$-a.s. and Proposition IV.2.3 shows that $t \to u_i(t,\cdot)$ is a continuous map from $(0,\infty)$ to $C_K(\mathbb{R})$ $\mathbb{P}_m^0$-a.s. It is now easy to see that $\mathbb{P}_m^0$-a.s. for all $0 < \delta \le t$,

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0} L_t^\varepsilon(X)(\phi) - L_\delta^\varepsilon(X)(\phi) \\
(IV.4.2) \qquad &= \lim_{\varepsilon\downarrow 0} \int_\delta^t \iint \phi(\frac{x_1+x_2}{2})p_\varepsilon(x_1 - x_2)u_1(s,x_1)u_2(s,x_2)dx_1dx_2ds \\
&= \int_\delta^t \int \phi(x)u_1(s,x)u_2(s,x)dxds.
\end{aligned}$$

Note also by the Domination Principle,

$$
\begin{aligned}
&\mathbb{P}^0_m\Big(L^\varepsilon_\delta(X)(1)+\int_0^\delta\int u_1(s,x)u_2(s,x)dxds\Big)\\
(IV.4.3)\qquad &\le \iint\int_0^\delta p_{2s+\varepsilon}(x_1-x_2)+p_{2s}(x_1-x_2)dsdm_1(x_1)dm_2(x_2)\\
&\le c\sqrt{\delta}m_1(1)m_2(1)\to 0 \text{ as } \delta\downarrow 0.
\end{aligned}
$$

(IV.4.1) now follows from (IV.4.2) and (IV.4.3). ∎

Recall that $\mathbb{P}^\varepsilon_m$ is the unique solution of $(CS)^{\varepsilon,\lambda}_m$ which is equivalent to $(GCS)^{\varepsilon,\lambda}_m$. In view of Remark IV.3.1(c) we may expect $\mathbb{P}^\varepsilon_m$ to converge to a solution of $(GCS)^\lambda_m$ as $\varepsilon\downarrow 0$.

Notation. $M_{FS}(\mathbb{R}^d)=\{m\in M_F(\mathbb{R}^d):\int_0^1 r^{1-d}\sup_x m(B(x,r))dr<\infty\}$.

If $m_1,m_2\in M_{FS}$, then an integration by parts shows that

$$\sup_{z_2}\int g_{d-2}(|z_1-z_2|)dm_1(z_1)<\infty$$

and so (m_1,m_2) satisfies the hypothesis (IC) of Theorem IV.3.8.

Theorem IV.4.2. (a) Assume $1\le d\le 3$ and $m\in(M_{FS})^2$.
(i) Then $\mathbb{P}^\varepsilon_m\overset{w}{\Rightarrow}\mathbb{P}_m$ on Ω^2, where $((X_t)_{t\ge 0},(\mathbb{P}_\nu)_{\nu\in M_{FS}})$ is an $(M_{FS})^2$-valued Borel $(\mathcal{F}^2_t)$-strong Markov process and $\mathbb{P}_m$ satisfies $(GCS)^\lambda_m$.
(ii) If, in addition, $\lambda_1=\lambda_2$. then $\mathbb{P}_m$ is the unique solution of $(GCS)^\lambda_m$.

(b) If $d=4$ or 5, $m\in(M_F(\mathbb{R}^d)-\{0\})^2$ satisfies the hypothesis (IC) of Theorem IV.3.2, and $\lambda\ne(0,0)$, then there is no solution to $(GCS)^\lambda_m$.

Discussion.(b) Theorem IV.3.8 shows that the existence of a collision local time for for any potential solutions of $(GCS)^\lambda_m$ is to be expected if $d\le 5$ and Theorem IV.3.2 suggests it will be nontrivial for $d\le 5$. These results may lead one to believe that nontrivial solutions exist for $d\le 5$. It turns out, however, that it is not the existence of collisions between a pair of independent super-Brownian motion that is germane to the existence of the solutions to (GCS). Rather it is the existence of collisions between a single Brownian path, B, and an independent super-Brownian motion, Z. If $G(B)=\{(t,B_t):t\ge 0\}$, then

$$(IV.4.4)\qquad \mathbb{P}(G(B)\cap G(Z)\ne\emptyset)>0 \text{ iff } d<4.$$

To see this for $d\ge 4$, recall from (IV.3.31) that $q^{\psi_d}(G(Z))<\infty$ a.s. We had $d>4$ there but the proof in Theorem 3.1 of Barlow, Evans and Perkins (1991) also goes through if $d=4$. This shows that $q^d(G(Z))=0$ if $d\ge 4$ and so (IV.4.4) is true by Theorem 1 of Taylor and Watson (1985) (i.e., the analogue of (IV.3.32) for $G(B)$). For $d\le 3$ one approach is to use a Tanaka formula to construct a nontrivial inhomogeneous additive functional of B which only increases on the set of times when $B(t)\in S(Z_t)$ (see Theorem 2.6 of Evans-Perkins (1998)). The construction requires a mild energy condition on the initial distributions of B and Z but the required result then holds for general initial conditions by Theorem III.2.2. Alternatively, a short direct proof using Theorem III.3.4 is given in Proposition 1.3 of Barlow and Perkins (1994).

To understand the relevance of (IV.4.4), we demonstrate its use in a heuristic proof of (b). Assume X satisfies $(GCS)_m^\lambda$ for $d=4$ or 5. Let $Z^i \geq X^i$ be a pair of dominating independent super-Brownian motions (from Proposition IV.1.4) and let H^i be the historical process associated with Z^i. The particle approximations in Example IV.1.1 suggest that X^1 is obtained from Z^1 by killing off some of the particles which collide with the X^2 population, and similarly for X^2. Use the notation of the Historical Cluster Representation (Theorem III.1.1) and let $\{y_1,\ldots,y_M\}$ be the finite support of $r_{t-\varepsilon}(H_t^1)$ for fixed $0<\varepsilon<t$. These are the ancestors at time $t-\varepsilon$ of the entire Z^1 population at time t. Which of these ancestors are still alive in the X^1 population at time $t-\varepsilon$? By Theorem III.3.1, y_i has law $\mathbb{E}(H^i_{t-\varepsilon}(\cdot))/m_1(1)$ and so is a Brownian path stopped at time $t-\varepsilon$ and is independent of Z^2. (IV.4.4) shows that $G(y_i)\cap G(Z^2)=\emptyset$ a.s. Therefore each y_i will not have encountered the smaller X^2 population up to time $t-\varepsilon$ and so must still be alive in the X^1 population. Let $\varepsilon\downarrow 0$ to see that the entire family tree of the population of Z^1 at time t never encounters Z^2 and hence X^2. This means that no particles have been killed off and so $Z_t^1=X_t^1$ a.s., and by symmetry, $Z_t^2=X_t^2$ a.s. These identities hold uniformly in t a.s. by continuity. The fact that $\mathbb{P}(L(Z)\neq 0)>0$ (Theorem IV.3.2) shows that Z does not satisfy $(GCS)_m^\lambda$ and so no solution can exist. In short, for $d=4$ or 5, the only collisions contributing to $L_t(Z)$ are between particles whose family trees die out immediately and so killing off these particles has no impact on the proposed competing species model.

The above proof is not hard to make rigorous if there is a historical process associated with X^i so that we can rigorously interpret the particle heuristics. To avoid this assumption, the proof given in Section 5 of Evans and Perkins (1994) instead uses the ideas underlying the Tanaka formula in the previous Section. The proof outlined above would also appear to apply more generally to any killing operation based on collisions of the two populations. In (GCS) we would replace $\lambda_i L_t(X)(\phi_i)$ with $A_t^i(\phi_i)$, where $A^i_\cdot$ is an increasing continuous M_F-valued process such that $S(A^i(dt,dx))\subset G(X^1)\cap G(X^2)$ a.s. The non-existence of solutions for $d=4$ or 5 in this more general setting is true (unpublished notes of Barlow, Evans and Perkins) but the 4-dimensional case is rather delicate.

(a) Tightness of $\{\mathbb{P}_m^\varepsilon\}$ is a simple exercise using the Domination Principle and Theorem IV.3.2. To show each limit point satisfies $(GCS)_m^\lambda$, a refinement of Theorem IV.3.8 is needed for $d\leq 3$ (see Theorem 5.10 of Barlow, Evans and Perkins (1991)). This refinement states that in (IV.3.10) the rate of convergence to 0 in probability is uniform in X satisfying (H_2). In the proof of (IV.3.10), the only step for which this additional uniformity requires $d\leq 3$ (and which requires some serious effort) is (IV.3.14). To handle this term we use Theorem III.3.4 to first bound the integrals with respect to $X_s^i(dx_i)$ at least if $d\leq 3$. If $\mathbb{P}_m^{\varepsilon_n}\overset{w}{\Rightarrow}$, use Skorohod's theorem to obtain solutions X^{ε_n} of $(CS)_m^{\varepsilon_n,\lambda}$ which converge a.s. to X, say, as $n\to\infty$. We now may let $n\to\infty$ in $(CS)_m^{\varepsilon_n,\lambda}$ to derive $(GCS)_m^\lambda$ for X–the above uniformity and a simple comparison of $L^{\varepsilon_n}(X^{\varepsilon_n})$ with $L^{\varepsilon_n,i}(X^{\varepsilon_n})$ (see Lemma 3.4 of Evans and Perkins (1994)) show that $L^{\varepsilon_n,i}(X^{\varepsilon_n})\to L(X)$ in probability as $n\to\infty$ and the other terms are easy to handle.

To proceed further seems to require considerable additional effort. The full convergence of the $\{P_m^\varepsilon\}$ to a nice strong Markov process is provided in Evans and Perkins (1998) (Theorems 1.6 and 8.2). Here we showed that each limit point has an associated pair of historical processes which satisfy a strong equation driven by a pair of independent historical Brownian motions whose supports carry a Poisson field of marks indicating potential killing locations. This strong equation has a solution which is unique, both pathwise and in law. (This general approach of using strong equations driven by historical processes will be used in another setting with greater attention to detail in the next Section.) This approach does show that the natural historical martingale problem associated with $(GCS)_m^\lambda$ is well-posed (Theorem 1.4 of Evans and Perkins (1998)). The uniqueness of solutions to $(GCS)_m^\lambda$ itself remains open in general as we do not know that any solution comes equipped with an associated historical process (from which we would be able to show it is the solution of the aforementioned strong equation). If $\lambda_1 = \lambda_2$, uniqueness of solutions to $(GCS)_m^\lambda$ was proved by Mytnik (1999) by a duality argument. Mytnik built a dual family of one-dimensional distributions (as opposed to a dual process) by means of an intricate and original Trotter product construction. One phase of the Trotter product requires solutions to a non-linear evolution equation with irregular initial data. As is often the case with duality arguments, it is non-robust and does not appear to handle the case where $\lambda_1 \neq \lambda_2$. It is somewhat disconcerting that after all of this effort the general question of uniqueness to our competing species model remains unresolved in general. I suspect the correct approach to these questions remains yet to be discovered and so was not tempted to provide a detailed description of the proofs here.

V. Spatial Interactions
1. A Strong Equation

We continue our study of measure-valued processes which behave locally like (A,γ,g)-DW superprocesses, i.e., where A, γ, and g may depend on the current state, X_t, of the process. In this Chapter we allow the generator governing the spatial motion, A, to depend on X_t. These results are taken from Perkins (1992), (1995). To simplify the exposition we set $\gamma=1$ and $g=0$, although as discussed below (in Section V.5) this restriction may be relaxed. Our approach may be used for a variety of dependencies of A_{X_t} on X_t but we focus on the case of state dependent diffusion processes. Let

$$\sigma : M_F(\mathbb{R}^d)\times\mathbb{R}^d\to\mathbb{R}^{d\times d},\quad b: M_F(\mathbb{R}^d)\times\mathbb{R}^d\to\mathbb{R}^d,\quad a=\sigma\sigma^*,$$

and set

$$A_\mu\phi(x)=\sum_i\sum_j a_{ij}(\mu,x)\phi_{ij}(x)+\sum_i b_i(\mu,x)\phi_i(x),\quad \text{for } \phi\in C_b^2(\mathbb{R}^d).$$

Here $a(\mu,x)$ and $b(\mu,x)$ are the diffusion matrix and drift of a particle at x in a population μ.

If $\mathrm{Lip}_1=\{\phi:\mathbb{R}^d\to\mathbb{R} : \|\phi\|_\infty\le 1,\ |\phi(x)-\phi(y)|\le\|x-y\|\ \forall x,y\in\mathbb{R}^d\}$ and μ, $\nu\in M_F(\mathbb{R}^d)$, the Vasershtein metric on $M_F(\mathbb{R}^d)$, introduced in Section II.7, is

$$d(\mu,\nu)=\sup\{|\mu(\phi)-\nu(\phi)| : \phi\in\mathrm{Lip}_1\}.$$

Recall that d is a complete metric on $M_F(\mathbb{R}^d)$ inducing the topology of weak convergence.

Our approach will be based on a fixed point argument and so we will need the following Lipschitz condition on b and σ:

Assume there is a non-decreasing function $L:\mathbb{R}_+\to\mathbb{R}_+$ such that

$$\text{(Lip)}\qquad \begin{aligned}&(a)\ \|\sigma(\mu,x)-\sigma(\mu',x')\|+\|b(\mu,x)-b(\mu',x')\|\\ &\qquad\le L(\mu(1)\vee\mu'(1))[d(\mu,\mu')+\|x-x'\|]\quad\forall \mu,\mu'\in M_F(\mathbb{R}^d),\ x,x'\in\mathbb{R}^d.\\ &(b)\ \sup_x\|\sigma(0,x)\|+\|b(0,x)\|<\infty.\end{aligned}$$

Remark V.1.1. (a) (Lip) easily implies that for some non-decreasing $C:\mathbb{R}_+\to\mathbb{R}_+$,

$$(B)\qquad \|\sigma(\mu,x)\|+\|b(\mu,x)\|\le C(\mu(1))\quad\forall\mu\in M_F(\mathbb{R}^d),\ x\in\mathbb{R}^d.$$

(b) The results of Sections V.1-V.4 remain valid without (Lip)(b) (see Section 5 of Perkins (1992)).

Exercise V.1.1. Prove that (Lip) holds in the following cases.
(a) $\sigma(\mu,x)=f(\mu(\phi_1),\dots,\mu(\phi_n),x)$, where ϕ_i are bounded Lipschitz functions on $\mathbb{R}^d$ and $f:\mathbb{R}^{n+d}\to\mathbb{R}^{d\times d}$ is Lipschitz continuous so that $\sup_x\|f(0,x)\|<\infty$.
(b) $b(\mu,x)=\sum_{k=1}^n\int b_k(x,x_1,\dots,x_k)d\mu(x_1)\dots d\mu(x_k)$
$\sigma(\mu,x)=\sum_{k=1}^n\int \sigma_k(x,x_1,\dots,x_k)d\mu(x_1)\dots d\mu(x_k)$,
where b_k and σ_k are bounded Lipschitz continuous functions taking values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$, respectively.

A special case of (b) would be $b(\mu,x)=\int b_1(x,x_1)d\mu(x_1)$ and $\sigma(\mu,x)=\int p_\varepsilon(x-x_1)d\mu(x_1)$. Here $b_1(x,x_1)\in\mathbb{R}^d$ models an attraction or repulsion between individuals at x and x_1, and particles diffuse at a greater rate if there are a number of other particles nearby.

To motivate our stochastic equation, consider the branching particle system in Section II.3 where $Y^\alpha\equiv B^\alpha$ are Brownian motions in $\mathbb{R}^d$, $X_0=m\in M_F(\mathbb{R}^d)$, and $\nu^n=\nu=\frac12\delta_0+\frac12\delta_2$. Recall from (II.8.3) that if $H_t^N=\frac1N\sum_{\alpha\sim t}\delta_{B^\alpha_{\cdot\wedge t}}$, then H^N converges weakly to a historical Brownian motion, H, with law $\mathbb{Q}_{0,m}$. Let $Z_0:\mathbb{R}^d\to\mathbb{R}^d$ be Borel. Now solve

$$(SE)_N(a)\ Z_t^\alpha=Z_0(B_0^\alpha)+\int_0^t\sigma(X_s^N,Z_s^\alpha)dB_s^\alpha+\int_0^t b(X_s^N,Z_s^\alpha)ds,\quad t<\frac{|\alpha|+1}{N}$$
$$(b)\ X_s^N=\frac1N\sum_{\beta\sim s}\delta_{Z_s^\beta}.$$

Such solutions are easy to construct in a pathwise unique manner on $[i/N,(i+1)/N)$ by induction on i. On $[i/N,(i+1)/N)$, we are solving a finite system of stochastic differential equations driven by $\{B_s^\alpha:\alpha\sim i/N,\ s\in[i/N,(i+1)/N)\}$ and with Lipschitz continuous coefficients. The latter uses

$$\begin{aligned} d\Big(\frac1N\sum_{\alpha\sim i/N}\delta_{x^\alpha},\frac1N\sum_{\alpha\sim i/N}\delta_{\hat x^\alpha}\Big)&\le N^{-1}\sum_{\alpha\sim i/N}\|x^\alpha-\hat x^\alpha\|\\ &\le (H^N_{i/N}(1))^{-1/2}N^{-1/2}\|x-\hat x\|_2,\end{aligned}$$

where $\|x-\hat x\|_2=\Big(\sum_{\alpha\sim i/N}\|x^\alpha-\hat x^\alpha\|^2\Big)^{1/2}$ and we have used Cauchy-Schwarz in the last inequality. This shows there is a pathwise solution to $(SE)_N$ on $[i/N,(i+1)/N)$. Now let the B^α's branch at $t=(i+1)/N$ and continue on $[(i+1)/N,(i+2)/N)$ with the new set of Brownian motions $\{B_s^\alpha:\alpha\sim(i+1)/N,\ s\in[(i+1)/N,(i+2)/N)\}$. These solutions are then pieced together to construct the $\{Z_t^\alpha:t<(|\alpha|+1)/N,\ \alpha\}$ in $(SE)_N$. If $N\to\infty$, we may expect $X^N\overset{w}{\Rightarrow}X$, where

$$(SE)\qquad (a)\ Z_t(\omega,y)=Z_0(y_0)+\int_0^t\sigma(X_s,Z_s)dy(s)+\int_0^t b(X_s,Z_s)ds$$
$$(b)\ X_t(\omega)(A)=\int 1(Z_t(\omega,y)\in A)H_t(\omega)(dy)\quad\forall A\in\mathcal{B}(\mathbb{R}^d)$$

The intuition here is that ω labels a tree of branching Brownian motions and y labels a branch on the tree. Then $Z_t(\omega,y)$ solves the sde along the branch y in the tree ω and $X_t(\omega)$ is the empirical distribution of these solutions. Our objective in this

Chapter is to give a careful interpretation of the stochastic integral in (SE)(a), prove that (SE) has a pathwise unique strong Markov solution and show that $X^N \overset{w}{\Rightarrow} X$.

2. Historical Brownian Motion

Throughout this Section we work in the setting of the historical process of Section II.8 where $(Y, P^x) \equiv (B, P^x)$ is d-dimensional Brownian motion. We adopt the notation given there with $E = \mathbb{R}^d$, but as B has continuous paths we replace $(D(E), \mathcal{D})$ with $(C, \mathcal{C})$, where $C = C(\mathbb{R}_+, \mathbb{R}^d)$ and $\mathcal{C}$ is its Borel σ-field. Let $\mathcal{C}_t = \sigma(y_s, s \le t)$ be its canonical filtration. If $Z : \mathbb{R}_+ \times C \to \mathbb{R}$, then

$$\begin{aligned} &Z \text{ is } (\mathcal{C}_t)\text{-predictable} \iff Z \text{ is } (\mathcal{C}_t)\text{-optional} \\ (V.2.1) \qquad &\qquad\qquad \iff Z \text{ is Borel measurable and } Z(t,y) = Z(t, y^t)\ \forall t \ge 0. \end{aligned}$$

This follows from Theorem IV.97 in Dellacherie and Meyer (1978) and the fact that the proofs given there remain valid if D is replaced by C. We will therefore identify Borel functions on $\hat{\mathbb{R}}^d = \{(t,y) \in \mathbb{R}_+ \times C : y = y^t\}$ with $(\mathcal{C}_t)$-predictable functions on $\mathbb{R}_+ \times C$. If $C^s = \{y \in C : y = y^s\}$ then this identification allows us to write the domain of the weak generator for the path-valued process W in (II.8.1) as

$$\begin{aligned} \mathcal{D}(\hat{A}) = \{\phi : \mathbb{R}_+ \times C \to \mathbb{R} : \ &\phi \text{ is bounded, continuous, and } (\mathcal{C}_t)\text{-predictable,} \\ &\text{and for some } \hat{A}_s\phi(y) \text{ with the same properties, } \phi(t,B) - \phi(s,B) \\ &- \int_s^t \hat{A}_r\phi(B)dr,\ t \ge s \text{ is a } (\mathcal{C}_t)\text{-martingale under } P_{s,y}\ \forall s \ge 0,\ y \in C^s\}. \end{aligned}$$

Recall here that $P_{s,y}$ is Wiener measure starting at time s with past history $y \in C^s$, and for $m \in M_F^s(C)$ (recall this means $y = y^s$ $m-$ a.s.) define $P_{s,m} = \int P_{s,y}m(dy)$.

For the rest of this Section assume $\tau \ge 0$ and $(K_t)_{t\ge\tau}$ satisfies $(HMP)_{\tau,K_\tau}$ (from Section II.8) on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge\tau}, \mathbb{P})$ with K_τ now possibly random with law ν, $\hat{\gamma} \equiv 1$, $\hat{g} \equiv 0$, and $\hat{A}$ equal to the generator of the path-valued Brownian motion described above. Assume that this probability space is complete, the filtration is right-continuous and $\mathcal{F}_\tau$ contains all the null sets. We also assume $\mathbb{E}(K_\tau(1)) < \infty$ so that $m(\cdot) = E(K_\tau(\cdot)) \in M_F^\tau(C)$ and we can still work with a martingale problem as opposed to a local martingale problem. Call such a process, K, an $(\mathcal{F}_t)$-historical Brownian motion starting at (τ, ν) (or (τ, m) if $\nu = \delta_m$). As in Theorem II.8.3, K is an $(\mathcal{F}_t)$-strong Markov process and has law $\mathbb{Q}_{\tau,\nu} = \int \mathbb{Q}_{\tau,K_0} d\nu(K_0)$. In this setting the superprocess property (II.8.5) becomes

$$(V.2.2) \qquad \mathbb{P}(K_t(\psi)) = P_{\tau,m}(\psi(B^t)) \quad \text{for } t \ge \tau,\ \psi \in b\mathcal{C}.$$

Note also that if $S \ge \tau$ is a finite valued $(\mathcal{C}_t)$-stopping time, then

$$(V.2.3) \qquad P_{\tau,m}(g|\mathcal{C}_S)(y) = P_{S,y^S}(g) \quad P_{\tau,m} - \text{a.a. } y\ \forall g \in b\mathcal{C}.$$

To see this write $g(y) = \tilde{g}(y^S, y(S+\cdot))$ and use the strong Markov property at time S.

Our main objective in this Section is the seemingly minor extension of $(HMP)_{\tau,K_\tau}$ presented in Proposition V.2.6 below, and the reader may want to skip ahead to this result and its Corollary V.2.7. The latter plays a key role in what follows. Note, however, that Proposition V.2.4 will also be used in our stochastic calculus on Brownian trees and the proof of Lemma V.2.2 illustrates a neat idea of Pat Fitzsimmons.

We first reduce the definition of $\mathcal{D}(\hat{A})$ to zero starting times.

Notation. $b\mathcal{P}$ is the space of bounded $\mathcal{C}_t$-predictable processes on $\mathbb{R}_+ \times C$.

Lemma V.2.1. $\phi \in \mathcal{D}(\hat{A})$ iff $\phi \in b\mathcal{P}$ is continuous and for some continuous $\hat{A}\phi \in b\mathcal{P}$,

$$n(t,y) = \phi(t,y) - \phi(0,y) - \int_0^t \hat{A}_r\phi(y)dr \text{ is a } (\mathcal{C}_t)\text{-martingale under } P^x \ \forall x \in R^d.$$

Proof. We need only show the above condition is sufficient for membership in $\mathcal{D}(\hat{A})$. Assume ϕ is as above and let $s > 0$. It suffices to show $n(t) - n(s),\ t \geq s$ is a $(\mathcal{C}_t)_{t\geq s}$-martingale under $P_{s,y}$ for every $y \in C^s$. Let $t \geq r \geq s$ and ψ be a bounded continuous $\mathcal{C}_r$-measurable mapping on C. We must show that

$$(V.2.4) \qquad P_{s,y}(n(t)\psi) = P_{s,y}(n(r)\psi) \quad \forall y \in C^s.$$

The left-hand side is

$$(V.2.5) \qquad P^0(n(t, y/s/(y(s)+B))\psi(y/s/(y(s)+B)))$$

and so is continuous in y by Dominated Convergence. The same is true of the right-hand side. It therefore suffices to establish (V.2.4) on a dense set of y in C^s. Next we claim that

$$(V.2.6) \qquad \text{the closed support of } P^{y_0}(B^s \in \cdot) \text{ is } \{y \in C^s : y(0) = y_0\}.$$

To see this first note that for every $\varepsilon,\ T > 0$, $P^{y_0}(\sup_{s\leq T}|B_s - y_0| < \varepsilon) > 0$ (e.g. by the explicit formula for the two-sided hitting time in Theorem 4.1.1 of Knight (1981)). Now use the classical Cameron-Martin-Girsanov formula to conclude that for any $\psi \in C(\mathbb{R}_+, \mathbb{R}^d)$, and $\varepsilon,\ T > 0$,

$$P^{y_0}(\sup_{s\leq t}|B_s - \int_0^s \psi(u)du - y_0| < \varepsilon) > 0.$$

The claim follows easily. It implies that (V.2.4) would follow from

$$P_{s,y^s}(n(t)\psi) = P_{s,y^s}(n(r)\psi) \quad P^{y_0} - \text{a.a. } y \ \text{ for all } y_0 \in \mathbb{R}^d.$$

By (V.2.3) this is equivalent to

$$P^{y_0}(n(t)\psi|\mathcal{C}_s)(y) = P^{y_0}(n(r)\psi|\mathcal{C}_s)(y) \quad P^{y_0} - \text{a.a. } y \text{ for all } y_0 \in \mathbb{R}^d.$$

This is immediate by first conditioning $n(t)\psi$ with respect to $\mathcal{C}_r$. ∎

Fitzsimmons (1988) showed how one can use Rost's theorem on balayage to establish sample path regularity of a general class of superprocesses. Although we have not needed this beautiful idea for our more restrictive setting, the next result illustrates its effectiveness.

Lemma V.2.2. Let $\phi,\ \psi : \mathbb{R}_+ \times C \to \mathbb{R}$ be $(\mathcal{C}_t)$-predictable maps such that for some fixed $T \geq \tau$, $\phi(t,y) = \psi(t,y)\ \forall \tau \leq t \leq T\ P_{\tau,m}$ – a.s. Then

$$\phi(t,y) = \psi(t,y)\ K_t - \text{a.a. } y \quad \forall \tau \leq t \leq T \quad \mathbb{P} - \text{a.s.}$$

Proof. Return to the canonical setting of historical paths, $(\Omega, \mathcal{F}_H, \mathcal{F}^H[\tau, t+], \mathbb{Q}_{\tau,\nu})$ of Section II.8, with $\hat{g} = 0$, $\hat{\gamma} = 1$, and P^x=Wiener measure. Recall the $\hat{E}$-valued diffusion $W_t = (\tau + t, Y^{\tau+t})$ with laws $\hat{P}_{\tau,y}$ and the W-superprocess

$$\hat{X}_t = \delta_{\tau+t} \times H_{\tau+t} \text{ with laws } \hat{\mathbb{P}}_{\tau,m}. \tag{V.2.7}$$

Note first that

(V.2.8)
if $g : [\tau, \infty) \times C \to \mathbb{R}_+$ is $(\mathcal{C}_t)_{t \geq \tau}$–predictable then $H_t(g_t)$ is $\mathcal{F}^H[\tau, t+]$–predictable.

To see this start with $g(t,y) = g_1(t) g_2(y^t)$, where g_1, g_2 are non-negative bounded continuous functions on $\mathbb{R}_+$ and C, respectively. Then (V.2.8) holds because $K_t(g_t)$ is a.s. continuous. A monotone class argument now proves (V.2.8) (recall (V.2.1)).

Let S be an $\mathcal{F}^H[\tau, t+])_{t \geq \tau}$-stopping time such that $\tau \leq S \leq T$ and let $\lambda > 0$. Then $\hat{S} = S - \tau$ is an $(\mathcal{F}_t^{\hat{X}})_{t \geq 0}$-stopping time. Define a finite measure μ on $\hat{E}$ by

$$\mu(g) = \mathbb{Q}_{\tau,\nu}(e^{-\lambda(S-\tau)} H_S(g_S)) = \hat{\mathbb{P}}_{\tau,\nu}(e^{-\lambda \hat{S}} \hat{X}_{\hat{S}}(g)),$$

where the second equality holds by Lemma II.8.1. Let $U_\lambda f$ be the λ-resolvent of W. If f is a non-negative function on $\hat{E}$, then the superprocess property ((II.8.5) and the display just before it) shows that

$$\begin{aligned}
\langle \delta_\tau \times m, U_\lambda f \rangle &= \hat{\mathbb{P}}_{\tau,\nu}\Big(\int_0^\infty e^{-\lambda t} \hat{X}_t(f)\, dt \Big) \\
&\geq \hat{\mathbb{P}}_{\tau,\nu}\Big(e^{-\lambda \hat{S}} \int_0^\infty e^{-\lambda t} \hat{X}_{t+\hat{S}}(f) dt \Big) \\
&= \hat{\mathbb{P}}_{\tau,\nu}\Big(e^{-\lambda \hat{S}} \hat{\mathbb{P}}_{\hat{X}_{\hat{S}}}\Big(\int_0^\infty e^{-\lambda t} \hat{X}_t(f) dt \Big)\Big) \\
&= \hat{\mathbb{P}}_{\tau,\nu}(e^{-\lambda \hat{S}} \hat{X}_{\hat{S}}(U_\lambda f)) = \langle \mu, U_\lambda f \rangle.
\end{aligned}$$

A theorem of Rost (1971) shows there is a randomized stopping time, V, on $C(\mathbb{R}_+, \hat{E}) \times [0,1]$ (i.e., V is jointly measurable and $\{y : V(y,u) \leq t\} \in \mathcal{C}_{t+}$ for all $u \in [0,1]$) such that for every non-negative Borel function g on $\hat{E}$,

$$\begin{aligned}
\mu(g) &= \int_0^1 \hat{P}_{\tau,m}(e^{-\lambda V(u)} g(W_{V(u)}))\, du \\
&\leq \int_0^1 P_{\tau,m}(g(\tau + V(u), Y^{\tau+V(u)}))\, du.
\end{aligned} \tag{V.2.9}$$

If $g(t,y) = |\phi(t,y) - \psi(t,y)|$, then the right-hand side of (V.2.9) is zero by hypothesis and so

$$H_S(|\phi(S) - \psi(S)|) = 0 \quad \text{a.s.}$$

The Section Theorem (Theorem IV.84 of Dellacherie and Meyer (1978)) and (V.2.8) then show that

$$H_t(|\phi(t)-\psi(t)|)=0 \quad \forall t\in[\tau,T] \quad \mathbb{Q}_{\tau,\nu}-\text{a.s.}$$

As K has law $\mathbb{Q}_{\tau,\nu}$, the result follows. ∎

Lemma V.2.3. Let $n:[\tau,\infty)\times C\to\mathbb{R}$ be a $(\mathcal{C}_t)_{t\ge\tau}$-predictable L^2-martingale under $P_{\tau,m}$. Then

$$K_t(n_t)=K_\tau(n_\tau)+\int_\tau^t\int n(s,y)dM(s,y) \quad \forall t\ge\tau \quad \mathbb{P}-\text{a.s.}$$

and is a continuous square integrable $(\mathcal{F}_t)$-martingale.

Proof. Let $N>\tau$. Then (V.2.3) and the Section Theorem imply that

$$(V.2.10) \qquad n(t,y)=P_{t,y^t}(n(N)) \quad \forall\tau\le t\le N \quad P_{\tau,m}-\text{a.s.}$$

Now let

$$\begin{aligned}S=\{X:C\to\mathbb{R}:X\in L^2(P_{\tau,m}),\ n^X(t,y)\equiv P_{t,y^t}(X)\text{ satisfies}\\ K_t(n_t^X)=K_\tau(n_\tau^X)+\int_\tau^t\int n^X(s,y)dM(s,y)\ \forall t\ge\tau\ \mathbb{P}-\text{a.s.}\}.\end{aligned}$$

Implicit in the above condition is that both sides are well-defined and finite. If $X\in C_b(C)$, then n^X is bounded and continuous on $\mathbb{R}_+\times C$ (recall (V.2.5)) and n^X is a continuous $(\mathcal{C}_t)$-martingale under P^x for all $x\in\mathbb{R}^d$ by (V.2.3). Lemma V.2.1 shows that $n^X\in\mathcal{D}(\hat{A})$ and $\hat{A}n^X=0$. $(HMP)_{\tau,K_\tau}$ therefore shows that $X\in S$.

Let $\{X_n\}\subset S$ and assume $X_n\xrightarrow{bp}X$. Then Dominated Convergence shows that $n^{X_n}\xrightarrow{bp}n^X$, $K_t(n_t^{X_n})\to K_t(n_t^X)$ $\forall t\ge\tau$, and (use (V.2.2))

$$\begin{aligned}&\mathbb{P}\Big(\int_\tau^t\int(n^{X_n}(s,y)-n^X(s,y))^2K_s(dy)ds\Big)\\&\quad=\int_\tau^t P_{\tau,m}((n^{X_n}(s,B^s)-n^X(s,B^s))^2)ds\to 0\text{ as }n\to\infty.\end{aligned}$$

Therefore we may let $n\to\infty$ in the equation showing $X^n\in S$ to conclude that $X\in S$. This and $C_b(C)\subset S$ show that $b\mathcal{C}\subset S$.

Let X be a non-negative function in $L^2(P_{\tau,m})$ and set $X_n=X\wedge n\in S$. Monotone Convergence shows that $n^{X_n}\uparrow n^X\le\infty$ pointwise and $K_t(n_t^{X_n})\uparrow K_t(n_t^X)$ for all $t\ge\tau$. (V.2.2) shows that

$$\begin{aligned}\mathbb{P}\Big(\int_\tau^t\int(n^{X_n}(s,y)-n^X(s,y))^2K_s(dy)ds\Big)&=\int_\tau^t P_{\tau,m}\Big((n^{X_n}(s,B^s)-n^X(s,B^s))^2\Big)ds\\&\le\int_\tau^t P_{\tau,m}((X_n-X)^2)ds\ \ \text{(by (V.2.3))}\\&\to 0\text{ as }n\to\infty.\end{aligned}$$

This allows us to conclude that $X \in S$, as above. It also shows that $\int_\tau^t \int n^X(s,y)dM(s,y)$ is a continuous L^2 martingale. In addition we have $\mathbb{P}(K_\tau(n_\tau^X)) = m(n_\tau^X) = P_{\tau,m}(X) < \infty$. All these results now extend to any X in $L^2(P_{\tau,m})$ by considering the positive and negative parts of X. Taking $X = n(N)$ we obtain the required result for $t \le N$, but with $\tilde{n}(t,y) = P_{t,y^t}(n(N))$ in place of $n(t,y)$. Now use (V.2.10) and Lemma V.2.2 to obtain the required result for $t \le N$ as none of the relevant quantities are changed for $t \le N$ off a $\mathbb{P}$-null set if we replace $\tilde{n}$ with n. Finally let $N \to \infty$. ∎

Definition. Let $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathcal{F}}_t) = (\Omega \times C, \mathcal{F} \times \mathcal{C}, \mathcal{F}_t \times \mathcal{C}_t)$ and let $\hat{\mathcal{F}}_t^*$ denote the universal completion of $\hat{\mathcal{F}}_t$. If T is a bounded $(\mathcal{F}_t)_{t\ge\tau}$-stopping time (write $T \in \mathcal{T}_b$ and note this means that $T \ge \tau$), the normalized Campbell measure associated with K_T is the probability $\hat{\mathbb{P}}_T$ on $(\hat{\Omega}, \hat{\mathcal{F}})$ given by

$$\hat{\mathbb{P}}_T(A \times B) = \mathbb{P}(1_A K_T(B))/m(1).$$

We denote sample points in $\hat{\Omega}$ by (ω, y). Therefore under $\hat{\mathbb{P}}_T$, ω has law $K_T(1)m(1)^{-1}d\mathbb{P}$ and given ω, y is then chosen according to $K_T(\cdot)/K_T(1)$. We will also consider $T \in \mathcal{T}_b$ as an $(\hat{\mathcal{F}}_t)$-stopping time and define $\hat{\mathcal{F}}_T$ accordingly.

Proposition V.2.4. (a) Assume $T \in \mathcal{T}_b$ and $\psi \in b\hat{\mathcal{F}}_T$, then

$$(V.2.11) \qquad K_t(\psi) = K_T(\psi) + \int_T^t \int \psi(y)dM(s,y) \quad \forall t \ge T \quad \mathbb{P}\text{-a.s.}$$

(b) Let $g : [\tau, \infty) \times \hat{\Omega}$ be $(\hat{\mathcal{F}}_t)$-predictable and bounded on $[\tau, N] \times \hat{\Omega}$ for all $N > \tau$. Then

$$(V.2.12) \quad \int \int_\tau^t g_s(\omega,y)ds K_t(dy) = \int_\tau^t \int \left[\int_\tau^s g_r(\omega,y)dr \right] dM(s,y) + \int_\tau^t K_s(g_s)ds \qquad \forall t \ge \tau \quad \text{a.s.}$$

Proof (a) Assume first T is constant and $\psi(\omega,y) = \psi_1(\omega)\psi_2(y^T)$ for $\psi_1 \in b\mathcal{F}_T$ and $\psi_2 : C \to \mathbb{R}$ bounded and continuous. Then

$$\phi(s,y) = P_{s,y^s}(\psi_2(B^T))$$

is a bounded predictable (C_t)-martingale under P^{y_0} for each $y_0 \in \mathbb{R}^d$ (use (V.2.2)) and is continuous on $\mathbb{R}_+ \times C$ (as in (V.2.5)). Lemma V.2.1 shows that $\phi \in \mathcal{D}(\hat{A})$ and $\hat{A}\phi = 0$. Therefore $(HMP)_{\tau,K_\tau}$ implies that for $t \ge T$,

$$\begin{aligned} K_t(\psi) = \psi_1 K_t(\phi_t) &= \psi_1 K_T(\psi_2) + \psi_1 \int_T^t \int \phi(s,y)dM(s,y) \\ &= K_T(\psi) + \int_T^t \int \psi(y)dM(s,y), \end{aligned}$$

because $\psi_1\phi(s) = \psi$ for $s \geq T$.

The proof now proceeds by a standard bootstrapping. The result clearly holds for ψ as above and T finite-valued and then for general T by the usual approximation of T by a decreasing sequence of finite-valued stopping times (the continuity of ψ_2 helps here). A monotone class argument now gives the result for any $\psi(\omega, y) = \tilde{\psi}(\omega, y^T)$, where $\tilde{\psi} \in b(\mathcal{F}_T \times \mathcal{C})$. We claim that any $\psi \in b\hat{\mathcal{F}}_T$ is of this form. For any $\psi \in b\hat{\mathcal{F}}_T$ there is an $(\hat{\mathcal{F}}_t)$-predictable process X so that $\psi = X(T)$ (Dellacherie and Meyer (1978), Theorem IV.67). It suffices to show that $X(T, \omega, y) = X(T, \omega, y^T)$ because we then prove the claim with $\tilde{\psi} = X(T)$. For this, first consider $X(t, \omega, y) = 1_{(s,u]}(t)1_A(\omega)1_B(y)$ for $u > s \geq \tau$, $A \in \mathcal{F}_s$ and $B \in \mathcal{C}_s$. Then the above claim is true because $1_B(y) = 1_B(y^s)$ and so on $\{s < T(\omega) \leq u\}$, $1_B(y) = 1_B(y^T(\omega))$. The aforementioned standard bootstrapping now gives the claim for any $(\hat{\mathcal{F}}_t)$-predictable X and so completes the proof of (a).

(b) First consider $g(s, \omega, y) = \phi(\omega, y)1_{(u,v]}(s)$ where $\phi \in b\hat{\mathcal{F}}_u$, $\tau \leq u < v$. Then $\mathbb{P}$-a.s. for $t \geq u$,

$$\begin{aligned}
\int \int_\tau^t g_s(\omega, y) ds K_t(dy) &= K_t(\phi)(t \wedge v - t \wedge u) \\
&= \int_u^t \int \phi(\omega, y)(s \wedge v - s \wedge u) dM(s, y) \\
&\quad + \int_\tau^t 1(u < s \leq v) K_s(\phi) ds \quad \text{(by (a) and integration} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{by parts)} \\
&= \int_\tau^t \int \left[\int_\tau^s g_r(\omega, y) dr \right] dM(s, y) + \int_\tau^t K_s(g_s) ds.
\end{aligned}$$

If $t < u$, the above equality holds because both sides are zero. The result therefore holds for linear combinations of the above functions, i.e., for $(\hat{\mathcal{F}}_t)$-simple g. Passing to the bounded pointwise closure we obtain the result for all $(\hat{\mathcal{F}}_t)$-predictable and bounded g. For g as in (b), we first get the result for $t \leq N$ by considering $g_{s\wedge N}$, and then for all t by letting $N \to \infty$. ∎

Remarks V.2.5. (a) If $g : [\tau, \infty) \times \hat{\Omega} \to \mathbb{R}$ is $(\hat{\mathcal{F}}_t^*)$-predictable and bounded, and μ is a σ-finite measure on $\hat{\Omega}$, then there are bounded $(\hat{\mathcal{F}}_t)$-predictable processes $g_1 \leq g \leq g_2$ such that $g_1(t) = g_2(t)$ $\forall t \geq \tau$ μ-a.e. This may be proved by starting with a simple g (i.e. $g(t, \omega, y) = \sum_{i=1}^{n} \phi_i(\omega, y)1_{(u_i, u_{i+1}]}(t) + \phi_0(\omega, y)1_{\{u_0 = t\}}$, where $\tau = u_0 < \ldots < u_{n+1} \leq \infty$, $\phi_i \in b\hat{\mathcal{F}}_{u_i}^*$) and using a monotone class theorem as on p. 134 of Dellacherie and Meyer (1978).

(b) If we take

$$\mu(A) = \mathbb{P}\left(\int_\tau^\infty \int 1_A(\omega, y) K_s(dy)ds\right)$$

in the above, then the right side of (V.2.12) is the same for g_1, g_2 and g. Here we have used the obvious extension of the stochastic integral with respect to M to $(\hat{\mathcal{F}}_t^*)$-predictable integrands. It follows from Proposition V.2.4 (b) that the left-hand side is the same for g_1 and g_2. By monotonicity it is the same for g and so (V.2.12) holds for $(\hat{\mathcal{F}}_t^*)$-predictable, bounded g. A straightforward truncation argument then gives it for $(\hat{\mathcal{F}}_t^*)$-predictable g satisfying

$$(V.2.13) \qquad \int_\tau^t \int \Big[\Big[\int_\tau^s |g_r| dr\Big]^2 + |g_s|\Big] K_s(dy)ds < \infty \quad \forall t > 0 \quad \mathbb{P}\text{-a.s.}$$

(c) In Proposition V.2.4 (a), if T is predictable, $\psi \in b\hat{\mathcal{F}}_{T-}^*$ (i.e. $(\hat{\mathcal{F}}^*_\cdot)_{T-}$), $g(t, \omega, y) = \psi(\omega, y) 1_{[T,\infty)}(t)$, (so g is $(\hat{\mathcal{F}}_t^*)$-predictable) and we take

$$\mu(A) = \mathbb{P}\left(\int 1_A(\omega, y) K_T(dy) + \int_T^\infty \int 1_A(\omega, y) K_s(dy)ds\right)$$

in (a), then $\psi_i = g_i(T, \omega, y) \in b\hat{\mathcal{F}}_{T-}$ (g_i as in (a)) and the right side of (V.2.11) is unchanged if ψ is replaced by ψ_i. As above, the inequality $\psi_1 \le \psi \le \psi_2$ shows the same is true of the left side. Therefore (V.2.11) remains valid if T is predictable and $\psi \in b\hat{\mathcal{F}}_{T-}^*$.

(d) If $\phi : [\tau, \infty) \times \hat{\Omega} \to \mathbb{R}$ is bounded and $(\hat{\mathcal{F}}_t^*)$-predictable, then $K_t(\phi_t)$ is $(\mathcal{F}_t)$-predictable. To see this, first note that this is clear from (c) (with $T = u$) if $\phi(t, \omega, y) = \psi(\omega, y) 1_{(u,v]}(t)$ for some $\tau \le u < v$ and $\psi \in b\hat{\mathcal{F}}_{u-}^*$. A monotone class argument now completes the proof (see Theorem IV.67 in Dellacherie and Meyer (1978)).

Here is the extension of $(HMP)_{\tau,K_\tau}$ we mentioned earlier.

Proposition V.2.6. Assume $\phi : [\tau, \infty) \times C \to \mathbb{R}$ is a $(\mathcal{C}_t)_{t\ge\tau}$-predictable map for which there is a $(\mathcal{C}_t)_{t\ge\tau}$-predictable map, $\bar{A}_{\tau,m}\phi = \bar{A}\phi$, such that

(*i*) $P_{\tau,m}\Big(\int_\tau^t \bar{A}\phi(s)^2\, ds\Big) < \infty \quad \forall\, t > \tau$

(*ii*) $n(t, y) = \phi(t, y) - \int_\tau^t \bar{A}\phi(s, y)ds,\ t \ge \tau$ is an L^2 $(\mathcal{C}_t)_{t\ge\tau}$-martingale under $P_{\tau,m}$.

Then

$$K_t(\phi_t) = K_\tau(\phi_\tau) + \int_\tau^t \int \phi(s, y)dM(s, y) + \int_\tau^t K_s(\bar{A}\phi_s)ds \quad \forall t \ge \tau \text{ a.s.}$$

The stochastic integral is a continuous L^2 martingale, $K_\tau(\phi_\tau)$ is square integrable, and $\mathbb{P}\Big(\int_\tau^t K_s(|\tilde{A}\phi_s|)ds\Big) < \infty$ for all $t \geq \tau$.

Proof. Note that (i), Cauchy-Schwarz and the superprocess property (V.2.2) show that (V.2.13) holds for $g = \bar{A}\phi$–in fact the expression there is integrable. Therefore we may use Remark V.2.5 (b) and Lemma V.2.3 to see that $\mathbb{P}$-a.s. for all $t \geq \tau$,

$$\begin{aligned}
K_t(\phi_t) &= K_t(n_t) + K_t\Big(\int_\tau^t \bar{A}\phi(s)\,ds\Big) \\
&= K_\tau(n_\tau) + \int_\tau^t\int n(s,y)dM(s,y) + \int_\tau^t\int\Big[\int_\tau^s \bar{A}\phi_r(y)dr\Big]dM(s,y) \\
&\qquad + \int_\tau^t K_s(\bar{A}\phi_s)ds \\
&= K_\tau(\phi_\tau) + \int_\tau^t\int \phi(s,y)dM(s,y) + \int_\tau^t K_s(\bar{A}\phi_s)ds.
\end{aligned}$$

Lemma V.2.3 shows that $K_\tau(\phi_\tau) = K_\tau(n_\tau)$ is square integrable and (V.2.2), (i) and (ii) show that

$$\mathbb{P}\Big(\int_\tau^t K_s(\phi_s^2)ds\Big) = \int_\tau^t P_{\tau,m}(\phi(s,B)^2)ds < \infty.$$

This shows the stochastic integral is an L^2 martingale and a similar argument shows that the drift term has integrable variation on bounded intervals. ∎

Notation. Let $\mathcal{D}(\bar{A}_{\tau,m})$ denote the class of ϕ considered in the above Proposition. Clearly $\mathcal{D}(\hat{A}) \subset \mathcal{D}(\bar{A}_{\tau,m})$ and $\bar{A}_{\tau,m}$ is an extension of $\hat{A}$ in that for any $\phi \in \mathcal{D}(\hat{A})$,

$$\int_\tau^t \bar{A}_{\tau,m}\phi(s,y)ds = \int_\tau^t \hat{A}\phi(s,y)ds \quad \forall t \geq \tau \quad P_{\tau,m} - \text{a.s.}$$

Exercise V.2.1. Assume $\tau = 0$ and K is an $(\mathcal{F}_t)$-historical Brownian motion starting at $(0,\nu)$–note we can treat K_0 as a finite measure on $\mathbb{R}^d$. Let $Z_0 : \mathbb{R}^d \to \mathbb{R}^d$ be a Borel map, $B \in \mathcal{B}(\mathbb{R}^d)$ and define $\tilde{Z}_0 : C \to C$ by

$$\tilde{Z}_0(y)(t) = y(t) - y(0) + Z_0(y(0)).$$

(a) If $\phi \in \mathcal{D}(\hat{A})$, show that $\tilde{\phi}(t,y) = \phi(t,\tilde{Z}_0(y))1_B(y(0)) \in \mathcal{D}(\bar{A}_{0,m})$ and

$$\bar{A}_{0,m}\tilde{\phi}(t,y) = \hat{A}\phi(t,\tilde{Z}_0(y))1_B(y(0)).$$

(b) Define $K'_t(F) = K_t(\{y : \tilde{Z}_0(y) \in F,\ y(0) \in B\})$ for $F \in \mathcal{C}$, and let ν' be the law of $K'_0 = K_0(Z_0^{-1}(A) \cap B)$ for $A \in \mathcal{B}(\mathbb{R}^d)$. Show that K' is an $(\mathcal{F}_t)$-historical Brownian motion starting at $(0,\nu')$ and therefore has law $\mathbb{Q}_{0,\nu'}$.

Hint. Use (a) to show that K' satisfies $(HMP)_{0,K'_\tau}$.

Corollary V.2.7. Let $T \in \mathcal{T}_b$ and assume $n(t),\ t \geq \tau$ is a $(\mathcal{C}_t)$-predictable square integrable martingale under $P_{\tau,m}$. Then $n(t \wedge T),\ t \geq \tau$ is an $(\hat{\mathcal{F}}_t)$-martingale under $\hat{\mathbb{P}}_T$.

Proof. Let $s \ge \tau$, $A \in \mathcal{F}_s$ and $B \in \mathcal{C}_s$. Define

$$\phi(t,y) = (n(t,y) - n(s,y))1_B(y)1(t \ge s).$$

Then $\phi \in D(\bar{A}_{\tau,m})$ and $\bar{A}_{\tau,m}\phi = 0$. Therefore $K_t(\phi_t)$ is an $(\mathcal{F}_t)$-martingale by Proposition V.2.6 and so

$$\begin{aligned}\hat{\mathbb{P}}_T((n(T) - n(s\wedge T))1_A(\omega)1_B(y)) &= \mathbb{P}(K_T(\phi_T)1_{A\cap\{T>s\}})m(1)^{-1}\\ &= \mathbb{P}(K_s(\phi_s)1_{A\cap\{T>s\}})m(1)^{-1} = 0,\end{aligned}$$

the last because $\phi_s = 0$. ∎

Example V.2.8. Recall $C_K^\infty(\mathbb{R}^k)$ is the set of infinitely differentiable functions on $\mathbb{R}^k$ with compact support and define

$$\begin{aligned}D_{fd} = \{\phi : \mathbb{R}_+ \times C \to \mathbb{R} : \phi(t,y) = \psi(y_{t_1\wedge t}, \dots, y_{t_n\wedge t}) \equiv \psi(\bar{y}_t),\\ 0 \le t_1 \le \dots \le t_n,\ \psi \in C_K^\infty(\mathbb{R}^{nd})\}.\end{aligned}$$

If ϕ is as above, let

$$\begin{aligned}&\phi_i(t,y) = \sum_{k=1}^n 1(t < t_k)\psi_{(k-1)d+i}(\bar{y}_t) \quad 1 \le i \le d,\\ &\phi_{ij}(t,y) = \sum_{k=1}^n\sum_{\ell=1}^n 1(t < t_k \wedge t_\ell)\psi_{(k-1)d+i,(\ell-1)+j}(\bar{y}_t) \quad 1 \le i,j \le d,\\ &\nabla\phi(t,y) = (\phi_1(t,y),\dots,\phi_d(t,y)), \text{ and } \bar{\Delta}\phi(t,y) = \sum_{i=1}^d \phi_{ii}(t,y).\end{aligned}$$

Itô's Lemma shows that for any $m \in M_F(C)$, $D_{fd} \subset D(\bar{A}_{\tau,m})$ and $A_{\tau,m}\phi = \frac{\bar{\Delta}\phi}{2}$ for $\phi \in D_{fd}$. In fact this remains true if C_K^∞ is replaced with C_b^2 in the above. Note that D_{fd} is not contained in $\mathcal{D}(\hat{A})$ because $\bar{\Delta}\phi(t,y)$ may be discontinuous in t.

Theorem V.2.9. Let $m \in M_F^\tau(C)$. An $(\mathcal{F}_t)_{t\ge\tau}$-adapted process $(K_t)_{t\ge\tau}$ with sample paths in $\Omega_H[\tau,\infty)$ is an $(\mathcal{F}_t)$-historical Brownian motion starting at (τ,m) iff for every $\phi \in D_{fd}$,

$$M_t(\phi) = K_t(\phi_t) - m(\phi_\tau) - \int_\tau^t K_s\Big(\frac{\bar{\Delta}\phi_s}{2}\Big)ds$$

is a continuous $(\mathcal{F}_t)$-local martingale such that $M_\tau(\phi) = 0$ and

$$\langle M(\phi)\rangle_t = \int_\tau^t K_s(\phi_s^2)ds \text{ for all } t \ge \tau \text{ a.s.}$$

Proof. The previous Example and Proposition V.2.6 show that an $(\mathcal{F}_t)$-historical Brownian motion does satisfy the above martingale problem.

The proof of the converse uses a generalization of the stochastic calculus developed in the next section for historical Brownian motion. It is proved in Theorem 1.3 of Perkins (1995)–see also Section 12.3.3 of Dawson (1993) for a different approach to a slightly different result. We will not use the uniqueness here although it plays an central role in the historical martingale problem treated in Perkins (1995) and discussed in Section 5 below. ∎

3. Stochastic Integration on Brownian Trees

Consider the question of defining the stochastic integral appearing in (SE)(a). We first need a probability measure on $\hat{\Omega}$ under which y is a Brownian motion. An infinite family of such probabilities is given below. We continue to work with the $(\mathcal{F}_t)$-historical Brownian motion, K_t, on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq\tau}, \mathbb{P})$ starting at (τ, ν) but now set $\tau = 0$ for convenience and so may view K_0 and its mean measure, m, as finite measures on $\mathbb{R}^d$.

Definition. Let $\bar{\Omega}' = (\Omega', \mathcal{G}, \mathcal{G}_t, \mathbb{Q})$ be a filtered space and T be a $(\mathcal{G}_t)$-stopping time. An $\mathbb{R}^d$-valued $(\mathcal{G}_t)$-adapted process, B_t, on $\bar{\Omega}'$ is a $(\mathcal{G}_t)$-Brownian motion stopped at T iff for $1 \leq i,j \leq d$, $B^i_t - B^i_0$ and $B^i_t B^j_t - \delta_{ij}(t \wedge T)$ are continuous $(\mathcal{G}_t)$-martingales.

If T is a constant time, (V.2.2) shows that under $\hat{\mathbb{P}}_T$, y is a Brownian motion stopped at T. The next result extends this to stopping times.

Proposition V.3.1. If $T \in \mathcal{T}_b$, then under $\hat{\mathbb{P}}_T$, y is a $(\hat{\mathcal{F}}_t)$-Brownian motion stopped at T.

Proof. Apply Corollary V.2.7 with $n(t,y) = y^i_t - y^i_0$ and $n(t,y) = y^i_t y^j_t - y^i_0 y^j_0 - \delta_{ij} t$. This gives the result because $y^i_t = y^i_{t\wedge T}$ $\hat{\mathbb{P}}_T$-a.s. ∎

Notation. $f \in D(n,d)$ iff $f : \mathbb{R}_+ \times \hat{\Omega} \to \mathbb{R}^{n\times d}$ is $(\hat{\mathcal{F}}^*_t)$-predictable and

$$\int_0^t \|f(s,\omega,y)\|^2 ds < \infty \ K_t - \text{a.a. } y \quad \forall t \geq 0 \quad \mathbb{P} - \text{a.s.}$$

Definition. If $X, Y : \mathbb{R}_+ \times \hat{\Omega} \to E$, we say $X = Y$ K-a.e. iff
$X(s,\omega,y) = Y(s,\omega,y)$ for all $s \leq t$ for K_t-a.a. y for all $t \geq 0$ $\mathbb{P}$-a.s.
If E is a metric space we say X is continuous K-a.e. iff $s \to X(s,\omega,y)$ is continuous on $[0,t]$ for K_t-a.a. y for all $t \geq 0$ $\mathbb{P}$-a.s.

If $T \in \mathcal{T}_b$ and $f \in D(n,d)$, then $\int_0^T \|f(s,\omega,y)\|^2 ds < \infty$ $\hat{\mathbb{P}}_T$-a.s. Therefore the classical stochastic integral $\int_0^t f(s,\omega,y) dy_s \equiv \hat{\mathbb{P}}_T - \int_0^t f(s,\omega,y) dy_s$ is uniquely defined up to $\hat{\mathbb{P}}_T$-null sets. The next result shows one can uniquely define a single process which represents these stochastic integrals for all T simultaneously.

Proposition V.3.2. (a) If $f \in D(n,d)$, there is an $\mathbb{R}^n$-valued $(\hat{\mathcal{F}}_t)$-predictable process $I(f,t,\omega,y)$ such that

$$(V.3.1)\ \ I(f, t\wedge T, \omega, y) = \hat{\mathbb{P}}_T - \int_0^t f(s,\omega,y) dy(s) \quad \forall t \geq 0 \quad \hat{\mathbb{P}}_T - \text{a.s. for all } T \in \mathcal{T}_b.$$

(b) If $I'(f)$ is an$(\hat{\mathcal{F}}^*_t)$-predictable process satisfying (V.3.1), then

$$I(f,s,\omega,y) = I'(f,s,\omega,y) \quad K - \text{a.e.}$$

(c) $I(f)$ is continuous K-a.e.

(d) (Dominated Convergence) For any $N>0$, if $f_k, f \in D(n,d)$ satisfy

$$\lim_{k\to\infty} \mathbb{P}\Big(K_N\Big(\int_0^N \|f_k(s)-f(s)\|^2 ds > \varepsilon\Big)\Big) = 0 \quad \forall\varepsilon>0,$$

then

$$\lim_{k\to\infty} \mathbb{P}\left(\sup_{t\le N} K_t\left(\sup_{s\le t}\|I(f_k,s,\omega,y)-I(f,s,\omega,y)\|>\varepsilon\right)\right)=0 \quad \forall\varepsilon>0.$$

(e) For any $S\in\mathcal{T}_b$ if $f_k, f\in D(n,d)$ satisfy

$$(V.3.2) \qquad \lim_{k\to\infty}\mathbb{P}\Big(K_S\Big(\int_0^S \|f_k(s)-f(s)\|^2 ds\Big)\Big)=0,$$

then

$$\sup_{t\le S} K_t(\sup_{s\le t}\|I(f_k,s)-I(f,s)\|^2)\xrightarrow{\mathbb{P}}0 \text{ as } k\to\infty.$$

Proof. To avoid factors of $m(1)^{-1}$ we will assume $m(1)=1$ throughout.
(b) Let

$$J(t,\omega)=\int \sup_{s\le t}\|I(f,s,\omega,y)-I'(f,s,\omega,y)\|\wedge 1\ K_t(dy).$$

Assume for the moment that J is $(\mathcal{F}_t)$-predictable, and let T be a bounded $(\mathcal{F}_t)$-predictable stopping time. Then

$$\mathbb{P}\left(J(T,\omega)\right)=\hat{\mathbb{P}}_T\left(\sup_{s\le T}\|I(f,s)-I'(f,s)\|\wedge 1\right)=0,$$

because under $\hat{\mathbb{P}}_T$, $I(f,s\wedge T)$ and $I'(f,s\wedge T)$ are both versions of $\hat{\mathbb{P}}_T-\int_0^s f(s)dy(s)$. By the Section Theorem we see that $J(t,\omega)=0$ $\forall t\ge 0$ $\mathbb{P}$-a.s., as required.

To prove J is $(\mathcal{F}_t)$-predictable, let $\phi(t,\omega,y)$ be the integrand in the definition of J. The projection of a $\mathcal{B}\times\hat{\mathcal{F}}_t^*$-measurable set onto Ω is $\hat{\mathcal{F}}_t^*$-measurable (Theorem III.13 of Dellacherie and Meyer (1978)) and so $\phi(t)$ is $(\hat{\mathcal{F}}_t^*)$-adapted. Therefore $\phi(t-)$ is $(\hat{\mathcal{F}}_t^*)$-predictable (being left-continuous) and hence so is

$$\phi(t)=\phi(t-)\vee\left(\|I(f,t,\omega,y)-I'(f,t,\omega,y)\|\wedge 1\right).$$

Remark V.2.5 (d) now shows that $J(t)=K_t(\phi(t))$ is $(\mathcal{F}_t)$-predictable.

(a), (c) For simplicity set $d=n=1$ in the rest of this proof (this only affects a few constants in what follows). If $f(s,\omega,y)=\sum_{i=1}^n f_i(\omega,y)1_{(u_i,u_{i+1}]}(s)$, where $f_i\in b\hat{\mathcal{F}}_{u_i}$ and $0=u_0<\ldots<u_{n+1}\le\infty$, (call f $(\hat{\mathcal{F}}_t)$-simple), then define

$$I(f,t,\omega,y)=\sum_{i=0}^n f_i(\omega,y)\left(y(t\wedge u_{i+1})-y(t\wedge u_i)\right).$$

This clearly satisfies the conclusions of (a) and (c).

Let $f \in D(1,1)$. As in the usual construction of the Itô integral we may choose a sequence of simple functions $\{f_k\}$ so that

$$(V.3.3) \qquad \hat{\mathbb{P}}_k\left(\int_0^k |f(s)-f_k(s)|^2\, ds \geq \frac{1}{4}2^{-3k}\right) \leq 2^{-k}.$$

Define

$$I(f,t,\omega,y) = \begin{cases} \lim_{k\to\infty} I(f_k,t,\omega,y) & \text{if it exists} \\ 0 & \text{otherwise} \end{cases}.$$

Clearly $I(f)$ is $(\hat{\mathcal{F}}_t)$-predictable.

Fix a bounded $(\mathcal{F}_t)$-stopping time T and let $m,\ell \geq n \geq T$. Use the fact that $I(f_m, t\wedge T,\omega,y)$ is a version of the $\hat{\mathbb{P}}_T$-Itô integral and standard properties of the latter to see that

$$\begin{aligned}
&\hat{\mathbb{P}}_T\left(\sup_{t\leq T}|I(f_m,t,\omega,y)-I(f_\ell,t,\omega,y)| > 2^{-n}\right)\\
&\leq \hat{\mathbb{P}}_T\left(\int_0^T |f_m(s)-f_\ell(s)|^2 ds \geq 2^{-3n}\right) + 2^{-n}\\
(V.3.4) \qquad &= \mathbb{P}\left(K_T\left(\int_0^T |f_m(s)-f_\ell(s)|^2 ds > 2^{-3n}\right)\right) + 2^{-n}\\
&= \mathbb{P}\left(K_n\left(\int_0^T |f_m(s)-f_\ell(s)|^2 ds > 2^{-3n}\right)\right) + 2^{-n}\\
&\qquad \text{(by Proposition V.2.4(a))}\\
&\leq \hat{\mathbb{P}}_n\left(\int_0^n |f_m(s)-f_\ell(s)|^2 ds > 2^{-3n}\right) + 2^{-n} \leq 3\cdot 2^{-n},
\end{aligned}$$

the last by (V.3.3) and an elementary inequality. This shows both that

$$(V.3.5) \qquad \sup_{t\leq T}|I(f_m,t,\omega,y)-I(f,t,\omega,y)| \to 0 \quad \text{as } m\to\infty \quad \hat{\mathbb{P}}_T\text{-a.s.}$$

and

$$\sup_{t\leq T}\left|\left(\hat{P}_T - \int_0^t f_m(s,\omega,y)dy(s)\right) - \left(\hat{P}_T - \int_0^t f(s,\omega,y)dy(s)\right)\right| \to 0 \quad \text{as } m\to\infty \quad \hat{\mathbb{P}}_T\text{-a.s.}$$

It follows that

$$\hat{\mathbb{P}}_T - \int_0^t f(s,\omega,y)dy(s) = I(f,t\wedge T,\omega,y) \quad \forall t\geq 0 \quad \hat{\mathbb{P}}_T\text{-a.s.}$$

because the left side is constant on $t \geq T$ $\hat{\mathbb{P}}_T$-a.s. This gives (a).

Set $m = n$ and let $\ell \to \infty$ in (V.3.4), and use (V.3.5) to conclude that

$$\sup_{T \in \mathcal{T}_b, T \leq m} \hat{\mathbb{P}}_T \left(\sup_{s \leq T} |I(f_m, s, \omega, y) - I(f, s, \omega, y)| > 2^{-m} \right) \leq 3 \cdot 2^{-n},$$

that is

$$\sup_{T \in \mathcal{T}_b, T \leq m} \mathbb{P} \left(K_T \Big(\sup_{s \leq T} |I(f_m, s, \omega, y) - I(f, s, \omega, y)| > 2^{-m} \Big) \right) \leq 3 \cdot 2^{-m}.$$

An application of the Section Theorem ($\tilde{J}(t, \omega) = K_t(\sup_{s \leq t} |I(f_m, s) - I(f, s)| > 2^{-m})$ is $(\mathcal{F}_t)$-predictable, as in (b)) gives

$$\mathbb{P} \left(\sup_{t \leq m} K_t \Big(\sup_{s \leq t} |I(f_m, s, \omega, y) - I(f, s, \omega, y)| > 2^{-m} \Big) \right) \leq 3 \cdot 2^{-m}.$$

Two successive applications of Borel-Cantelli show that

$$(V.3.6) \qquad \lim_{m \to \infty} \sup_{s \leq t} |I(f_m, s, \omega, y) - I(f, s, \omega, y)| = 0 \quad K_t\text{-a.a. y} \quad \forall t \geq 0 \quad \text{a.s.}$$

This certainly implies (c).

(d) Fix $N > 0$ and assume $\{f_k\}$, f satisfy the hypotheses of (d). Argue exactly as in (V.3.4) with $\varepsilon > 0$ in place of 2^{-n}, but now use Remark V.2.5 (c) in place of Proposition V.2.4 (a) and take the sup over $(\mathcal{F}_t)$-predictable times to conclude

$$\begin{aligned} &\sup_{T \leq N, T \text{ predictable}} \hat{\mathbb{P}}_T \left(\sup_{t \leq T} |I(f_k, t, \omega, y) - I(f, t, \omega, y)| > \varepsilon \right) \\ &\qquad \leq \hat{\mathbb{P}}_N \left(\int_0^N |f_k - f(s, \omega, y)|^2 ds > \varepsilon^3 \right) + \varepsilon. \end{aligned}$$

The first term on the right-hand side approaches 0 as $k \to \infty$. As ε is arbitrary, the same is true for the left-hand side. As in (b),

$$(t, \omega) \to K_t \left(\sup_{s \leq t} |I(f_k, s, \omega, y) - I(f, s, \omega, y)| > \varepsilon \right)$$

is $(\mathcal{F}_t)$-predictable and the Section Theorem implies

$$\sup_{t \leq N} K_t \left(\sup_{s \leq t} |I(f_k, s, \omega, y) - I(f, s, \omega, y)| > \varepsilon \right) \xrightarrow{\mathbb{P}} 0 \quad \text{as} \quad k \to \infty \quad \forall \varepsilon > 0.$$

The random variables on the left are all bounded by $\sup_{t \leq N} K_t(1) \in L^1$ and so also converge in L^1 by Dominated Convergence for all $\varepsilon > 0$. This is the required conclusion in (d).

(e) Let $S, T \in \mathcal{T}_b$ with $T \le S$. Doob's maximal L^2 inequality shows that

$$\begin{aligned}\mathbb{P}\Big(K_T\Big(\sup_{s\le T}(I(f,s)-I(f_k,s))^2\Big)\Big)) &\le c\hat{\mathbb{P}}_T\Big(\int_0^T (f(s)-f_k(s))^2\,ds\Big)\\ &= c\mathbb{P}\Big(K_S\Big(\int_0^T (f(s)-f_k(s))^2\,ds\Big)\Big)\\ &\le c\mathbb{P}\Big(K_S\Big(\int_0^S (f(s)-f_k(s))^2\,ds\Big)\Big).\end{aligned}$$

Remark V.2.5 (c) was used in the second line. Therefore (V.3.2) implies that

$$(V.3.7) \qquad \lim_{k\to\infty}\sup_{T\le S, T\in\mathcal{T}_b}\mathbb{P}\Big(K_T\Big(\sup_{s\le T}(I(f,s)-I(f_k,s))^2\Big)\Big)=0.$$

As in the proof of (b), $(t,\omega)\to K_t(\sup_{s\le t}(I(f,s)-I(f_k,s))^2)$ is $(\mathcal{F}_t)$-predictable. A simple application of the Section Theorem shows that

$$\begin{aligned}&\mathbb{P}\Big(\sup_{t\le S}K_t\Big(\sup_{s\le t}(I(f,s)-I(f_k,s))^2\Big)>\varepsilon\Big)\\ &\quad=\sup_{T\le S,T\in\mathcal{T}_b}\mathbb{P}\Big(K_T\Big(\sup_{s\le T}(I(f,s)-I(f_k,s))^2\Big)>\varepsilon\Big),\end{aligned}$$

which approaches zero as $k\to\infty$ by (V.3.7). ∎

Lemma V.3.3. Let $g:\mathbb{R}_+\times\hat{\Omega}\to\mathbb{R}_+$ be $(\hat{\mathcal{F}}_t^*)$-predictable and $S\in\mathcal{T}_b$.
(a) $\mathbb{P}\Big(K_S\Big(\int_0^S g_s ds\Big)\Big)=\mathbb{P}\Big(\int_0^S K_s(g_s)ds\Big)$.
(b) $\mathbb{P}\Big(\sup_{t\le S}K_t\Big(\int_0^t g_s ds\Big)>\varepsilon\Big)\le\varepsilon^{-1}\mathbb{P}\Big(\int_0^S K_s(g_s)ds\Big)$ for all $\varepsilon>0$.

Proof. (a) By Monotone Convergence it suffices to consider g bounded. This case is then immediate from Remark V.2.5(b).
(b) From Remark V.2.5 (d) we see that $K_t\Big(\int_0^t g_s\,ds\Big)\le\infty$ is $(\mathcal{F}_t)$-predictable. By the Section Theorem,

$$\begin{aligned}\mathbb{P}\Big(\sup_{t\le S}K_t\Big(\int_0^t g_s ds\Big)>\varepsilon\Big) &= \sup_{T\le S,T\in\mathcal{T}_b}\mathbb{P}\Big(K_T\Big(\int_0^T g_s ds\Big)>\varepsilon\Big)\\ &\le \sup_{T\le S,T\in\mathcal{T}_b}\varepsilon^{-1}\mathbb{P}\Big(K_T\Big(\int_0^T g_s ds\Big)\Big)\\ &=\varepsilon^{-1}\mathbb{P}\Big(\int_0^S K_s(g_s)ds\Big)\quad\text{(by (a)).}\end{aligned}$$ ∎

Notation. If $X(t)=(X_1(t),\dots,X_n(t))$ is an $\mathbb{R}^n$-valued process on $(\hat{\Omega},\hat{\mathcal{F}})$ and $\mu\in M_F(\mathbb{R}^d)$, let $\mu(X_t)=(\mu(X_1(t)),\dots,\mu(X_n(t)))$ and

$$\int_0^t\int X(s)dM(s,y)=\Big(\int_0^t\int X_1(s)dM(s,y),\dots,\int_0^t\int X_n(s)dM(s,y)\Big),$$

whenever these integrals are defined. We also write $\int_0^t f(s,\omega,y)dy(s)$ for $I(f,t,\omega,y)$ when $f \in D(n,d)$ and point out that dependence on K is suppressed in either notation.

Proposition V.3.4. If $f \in D(n,d)$ is bounded, then

$$(V.3.8) \qquad K_t(I(f,t)) = \int_0^t \int I(f,s)dM(s,y) \quad \forall t \geq 0 \quad \mathbb{P}-\text{a.s.},$$

and the above is a continuous L^2 $(\mathcal{F}_t)$-martingale.

Proof. To simplify the notation take $n = d = 1$ and $m(1) = 1$. Assume first that

$$(V.3.9) \quad f(s,\omega,y) = \phi_1(\omega)\phi_2(y)1(u < s \leq v), \quad \phi_1 \in b\mathcal{F}_u,\ \phi_2 \in b\mathcal{C}_u,\ 0 \leq u < v.$$

If $n(t,y) = \phi_2(y)(y(t \wedge v) - y(t \wedge u))$, then $(n(t), \mathcal{C}_t)$ is an L^2-martingale under $P_{0,m}$ and $I(f,t) = \phi_1(\omega)n(t,y)$. Lemma V.2.3 shows that $\mathbb{P}$-a.s. for all $t \geq 0$,

$$\begin{aligned} K_t(I(f,t)) &= \phi_1(\omega)K_t(n_t) \\ &= \phi_1(\omega)\int_0^t \int n(s,y)dM(s,y) \\ &= \int_0^t \int I(f,s)dM(s,y). \end{aligned}$$

In the last line we can take ϕ_1 through the stochastic integral since $\phi_1 \in b\mathcal{F}_u$ and $n(s)$ vanishes for $s \leq u$.

Suppose (V.3.8) holds for a sequence of $(\hat{\mathcal{F}}_t)$-predictable processes f_k and f is an $(\hat{\mathcal{F}}_t^*)$-predictable process such that $\sup_k \|f_k\|_\infty \vee \|f\|_\infty < \infty$ and

$$(V.3.10) \qquad \lim_{k\to\infty} \mathbb{P}\Big(\int \int_0^N (f_k(s) - f(s))^2 ds K_N(dy)\Big) = 0 \quad \forall N \in \mathbb{N}.$$

We claim (V.3.8) also holds for f. If $N \in \mathbb{N}$, then use the fact that under $\hat{\mathbb{P}}_s$, $I(f)$ is an ordinary Itô integral to conclude

$$\begin{aligned} \mathbb{P}\Big(\int_0^N \int (I(f,s) - I(f_k,s))^2 K_s(dy)ds\Big) &= \int_0^N \hat{\mathbb{P}}_s((I(f,s) - I(f_k,s))^2)ds \\ &= \int_0^N \hat{\mathbb{P}}_s\Big(\int_0^s (f(r) - f_k(r))^2 dr\Big)ds \\ &= \int_0^N \mathbb{P}\Big(K_N\Big(\int_0^s (f(r) - f_k(r))^2 dr\Big)\Big)ds, \end{aligned}$$

where Remark V.2.5 (c) is used in the last line. This approaches zero as $k \to \infty$ by (V.3.10). Therefore $\int_0^t \int I(f,s)dM(s,y)$ is a continuous L^2 martingale and

$$(V.3.11) \qquad \lim_{k\to\infty} \Big\| \sup_{t\leq N} \Big| \int_0^t \int I(f_k,s) - I(f,s)dM(s,y) \Big| \Big\|_2 = 0 \quad \forall N \in \mathbb{N}.$$

Proposition V.3.2 (e) with $S = N$ and (V.3.10) imply

$$(V.3.12) \qquad \sup_{t \le N} K_t(|I(f_k,t) - I(f,t)|) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty \quad \forall N \in \mathbb{N}.$$

Now let $k \to \infty$ in (V.3.8) (for f_k) to see that it also holds for f.

We may now pass to the bounded pointwise closure of the linear span of the class of f satisfying (V.3.9) to see that the result holds for all bounded $(\hat{\mathcal{F}}_t)$-predictable f. If f is bounded and $(\hat{\mathcal{F}}^*_t)$-predictable, there is a bounded $(\hat{\mathcal{F}}_t)$-predictable $\tilde{f}$ so that

$$\mathbb{P}\Big(\int \int_0^N (f(s) - \tilde{f}(s))^2 ds K_N(dy)\Big) = 0 \quad \forall N \in \mathbb{N}$$

(see Remark V.2.5 (a)) and so the the result holds for f by taking $f_k = \tilde{f}$ in the above limiting argument. ∎

Theorem V.3.5 (Itô's Lemma). Let Z_0 be $\mathbb{R}^n$-valued and $\hat{\mathcal{F}}^*_0$-measurable, $f \in D(n,d)$, g be an $\mathbb{R}^n$-valued $(\hat{\mathcal{F}}^*_t)$-predictable process and $\psi \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^n)$. Assume

$$(V.3.13) \qquad \int_0^t K_s(\|f_s\|^2 + \|g_s\|) ds < \infty \quad \forall t > 0 \text{ a.s.},$$

and let

$$(V.3.14) \qquad Z_t(\omega,y) = Z_0(\omega,y) + \int_0^t f(s,\omega,y) dy_s + \int_0^t g(s,\omega,y) ds.$$

If $\nabla\psi$ and ψ_{ij} denote the gradient and second order partials in the spatial variables, then

$$\int \psi(t,Z_t) K_t(dy) = \int \psi(0,Z_0) dK_0(y) + \int_0^t \int \psi(s,Z_s) dM(s,y)$$
$$(V.3.15) \qquad + \int_0^t K_s\Big(\frac{\partial \psi}{\partial s}(s,Z_s) + \nabla\psi(s,Z_s)\cdot g_s + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \psi_{ij}(s,Z_s)(ff^*)_{ij}(s)\Big) ds.$$

The second term on the right is an L^2 martingale and the last term on the right has paths with finite variation on compact intervals a.s. In particular $\int \psi(t,Z_t)K_t(dy)$ is a continuous $(\mathcal{F}_t)$-semimartingale.

Proof. Assume first that $\|f\|$ and $\|g\|$ are bounded. Let $T \in \mathcal{T}_b$ and $\mathbf{Z}_t = (t, Z_t)$. Itô's Lemma shows that $\hat{\mathbb{P}}_T$-a.s.

$$\psi(\mathbf{Z}_{t\wedge T}) = \psi(\mathbf{Z}_0) + \hat{\mathbb{P}}_T - \int_0^{t\wedge T} \nabla\psi(\mathbf{Z}_s) f(s)\cdot dy(s)$$
$$+ \int_0^{t\wedge T} \frac{\partial\psi}{\partial s}(\mathbf{Z}_s) + \nabla\psi(\mathbf{Z}_s)\cdot g(s) + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n \psi_{ij}(\mathbf{Z}_s)(ff^*)_{ij}(s) ds \ \forall t \ge 0.$$

Let $\tilde{b}(s)$ denote the integrand in the last term. This shows that

$$\tilde{I}(t) = \psi(\mathbf{Z}_t) - \psi(\mathbf{Z}_0) - \int_0^t \tilde{b}(s)ds$$

is a $(\hat{\mathcal{F}}_t^*)$-predictable process satisfying (V.3.1) with $\nabla\psi(\mathbf{Z}_s)f(s)$ in place of $f(s)$. Proposition V.3.2 therefore implies that

$$(V.3.16) \qquad \psi(\mathbf{Z}_t) = \psi(\mathbf{Z}_0) + I(\nabla\psi(\mathbf{Z})f, t) + \int_0^t \tilde{b}(s)ds \quad K - \text{a.e.}$$

Since $\|\tilde{b}\|$ and $\|\nabla\psi(\mathbf{Z})f\|$ are bounded and $(\hat{\mathcal{F}}_T^*)$-predictable we may apply Proposition V.3.4 and Remarks V.2.5 to see that $\mathbb{P}$-a.s. for all $t \geq 0$,

$$\begin{aligned}
&\int \psi(\mathbf{Z}_t)K_t(dy) \\
&\quad = \int \psi(\mathbf{Z}_0)K_0(dy) + \int_0^t \int \psi(\mathbf{Z}_0)dM(s,y) + \int_0^t \int I(\nabla\psi(\mathbf{Z})f, s)dM(s,y) \\
&\qquad + \int_0^t \int \Big[\int_0^s \tilde{b}(r)dr\Big]dM(s,y) + \int_0^t K_s(\tilde{b}(s))ds \\
&\quad = \int \psi(\mathbf{Z}_0)K_0(dy) + \int_0^t \int \psi(\mathbf{Z}_s)dM(s,y) + \int_0^t K_s(\tilde{b}(s))ds.
\end{aligned}$$

In the last line we used (V.3.16) and the fact that this implies the stochastic integrals of both sides of (V.3.16) with respect to M coincide. This completes the proof in this case.

Assume now that f, g satisfy (V.3.13). By truncating we may choose bounded $(\mathcal{F}_t^*)$-predictable f^k, g^k such that $f^k \to f$ and $g^k \to g$ pointwise, $\|f^k\| \leq \|f\|$ and $\|g^k\| \leq \|g\|$ pointwise, and therefore

$$(V.3.17) \qquad \lim_{k\to\infty} \int_0^t K_s(\|f_s^k - f_s\|^2 + \|g_s^k - g_s\|)ds = 0 \quad \forall t \geq 0 \text{ a.s.}$$

By (V.3.13) we may choose $S_n \in \mathcal{T}_b$ satisfying $S_n \uparrow \infty$ a.s. and

$$\int_0^{S_n} K_s(\|f_s\|^2 + \|g_s\|)ds \leq n \quad \text{a.s.}$$

Define Z^k as in (V.3.14) but with (f^k, g^k) in place of (f, g). Note that Lemma V.3.3(b) shows that $\sup_{t\leq S_n} K_t\Big(\int_0^t \|g_s\|ds\Big) < \infty$ for all n a.s. and so $Z_t(\omega, y)$ is well-defined K_t-a.a. y for all $t \geq 0$ a.s. The same result shows that

$$(V.3.18) \qquad \sup_{t\leq S_n} K_t\Big(\int_0^t \|g_s^k - g_s\|ds\Big) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty \quad \forall n.$$

A similar application of Proposition V.3.2 (e) and Lemma V.3.3(a) and gives

$$(V.3.19) \qquad \sup_{t\leq S_n} K_t(\sup_{s\leq t} \|I(f^k, s) - I(f, s)\|^2) \xrightarrow{\mathbb{P}} 0 \text{ as } k \to \infty \quad \forall n.$$

(V.3.18) and (V.3.19) imply

$$(V.3.20) \qquad \sup_{t\le T} K_t(\sup_{s\le t}\|Z^k(s)-Z(s)\|)\xrightarrow{\mathbb{P}}0 \text{ as } k\to\infty \quad \forall T>0.$$

We have already verified (V.3.15) with (Z^k, f^k, g^k) in place of (Z, f, g). The boundedness of ψ and its derivatives, together with (V.3.17) and (V.3.20), allow us to let $k\to\infty$ in this equation and use Dominated Convergence to derive the required result. (Clearly (V.3.13) implies the "drift" term has bounded variation on bounded time intervals a.s.) ∎

Corollary V.3.6. If f, g, and Z are as in Theorem V.3.5, then

$$X_t(A)=\int 1(Z_t(\omega,y)\in A)K_t(dy)$$

defines an a.s. continuous $(\mathcal{F}_t)$-predictable $M_F(\mathbb{R}^n)$-valued process.

Proof. Take $\psi(x)=e^{iu\cdot x}$, $u\in\mathbb{Q}^n$ to see that $\int e^{iu\cdot x}dX_t(x)$ is continuous in $t\in[0,\infty)$ for all $u\in\mathbb{Q}^n$ a.s. Lévy's Continuity Theorem completes the proof of continuity. Remark V.2.5 (d) shows that X is $(\mathcal{F}_t)$-predictable. ∎

4. Pathwise Existence and Uniqueness

As in the last Section we assume $(K_t, t\ge 0)$ is an $(\mathcal{F}_t)$-historical Brownian motion on $\bar\Omega=(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$ starting at $(0,\nu)$ with $\mathbb{E}(K_0(\cdot))=m(\cdot)\in M_F(\mathbb{R}^d)$. Therefore K has law $\mathbb{Q}_{0,\nu}$ on $(\Omega_X,\mathcal{F}_X)$. Recall that σ, b satisfy (Lip) and $Z_0:\mathbb{R}^d\to\mathbb{R}^d$ is a Borel map. If $\int_0^t\sigma(X_s,Z_s)dy(s)$ is the stochastic integral introduced in Proposition V.3.1, here is the precise interpretation of (SE):

$$(SE)_{Z_0,K}\ (a)\ Z(t,\omega,y)=Z_0(y_0)+\int_0^t\sigma(X_s,Z_s)dy(s)+\int_0^t b(X_s,Z_s)ds \quad K-\text{a.e.}$$

$$(b)\ X_t(\omega)(A)=\int 1(Z_t(\omega,y)\in A)K_t(\omega)(dy)\quad \forall A\in\mathcal{B}(\mathbb{R}^d)\ \forall t\ge 0 \text{ a.s.}$$

(X,Z) is a solution of $(SE)_{Z_0,K}$ iff Z is a $(\hat{\mathcal{F}}_t^*)$-predictable $\mathbb{R}^d$-valued process, and X is an $(\mathcal{F}_t)$-predictable $M_F(\mathbb{R}^d)$-valued process such that $(SE)_{Z_0,K}$ holds. Let $\bar{\mathcal{H}}_t$ denote the usual enlargement of $\mathcal{F}^H[0,t+]$ with $\mathbb{Q}_{0,\nu}$-null sets.

Theorem V.4.1.

(a) There is a pathwise unique solution (X,Z) to $(SE)_{Z_0,K}$. More precisely X is unique up to $\mathbb{P}$-evanescent sets and Z is unique K-a.e. Moreover $t\mapsto X_t$ is a.s. continuous in t.

(b) There are $(\bar{\mathcal{H}}_t)$-predictable and $(\hat{\bar{\mathcal{H}}}{}_t^*)$-predictable maps, $\tilde X:\mathbb{R}_+\times\Omega_H\to M_F(\mathbb{R}^d)$ and $\tilde Z:\mathbb{R}_+\times\hat\Omega_H\to\mathbb{R}^d$, respectively, which depend only on (Z_0,ν), and are such that that $(X(t,\omega),Z(t,\omega,y))=(\tilde X(t,K(\omega)),\tilde Z(t,K(\omega),y))$ defines the unique solution of $(SE)_{Z_0,K}$.

(c) There is a continuous map $X_0\to\mathbb{P}'_{X_0}$ from $M_F(\mathbb{R}^d)$ to $M_1(\Omega_X)$, such that if (X,Z) is a solution of $(SE)_{Z_0,K}$ on some filtered space $\bar\Omega$, then

$$(V.4.1) \qquad \mathbb{P}(X\in\cdot)=\int\mathbb{P}'_{X_0(\omega)}d\mathbb{P}(\omega).$$

(d) If T is an a.s. finite $(\mathcal{F}_t)$-stopping time, then

$$\mathbb{P}(X(T+\cdot)\in A|\mathcal{F}_T)(\omega)=\mathbb{P}'_{X_T(\omega)}(A)\quad \mathbb{P}-\text{a.s. for all } A\in\mathcal{F}_X.$$

For the uniqueness and continuity of $\mathbb{P}'_{X_0}$ we will need to prove the following stability result. Recall that d is the Vasershtein metric on $M_F(\mathbb{R}^d)$.

Theorem V.4.2. Let $K^1\le K^2\equiv K$ be $(\mathcal{F}_t)$-historical Brownian motions with $\mathbb{E}(K_0^i(\cdot))=m_i(\cdot)\in M_F(\mathbb{R}^d)$, and let $Z_0^i:\mathbb{R}^d\to\mathbb{R}^d$ be Borel maps, $i=1,2$. Set $T_N=\inf\{t:K_t(1)\ge N\}\wedge N$. There are universal constants $\{c_N, N\in\mathbb{N}\}$ so that if (X^i,Z^i) is the unique solution of $(SE)_{Z_0^i,K^i}$ then

$$\mathbb{P}\Big(\int_0^{T_N}\sup_{u\le s} d(X^1(u),X^2(u))^2ds\Big)\le c_N\Big[\int\|Z_0^1-Z_0^2\|^2\wedge 1dm_2+m_2(1)-m_1(1)\Big].$$

Here are the metric spaces we will use in our fixed point argument:

$$\begin{aligned}
S_1&=\{X:\mathbb{R}_+\times\Omega\to M_F(\mathbb{R}^d): X \text{ is } (\mathcal{F}_t)-\text{predictable and a.s. continuous },\\
&\qquad\qquad\qquad\qquad\qquad\qquad X_t(1)\le N\ \forall t<T_N\ \forall N\in\mathbb{N}\},\\
S_2&=\{Z:\mathbb{R}_+\times\hat\Omega\to\mathbb{R}^d: Z \text{ is } (\hat{\mathcal{F}}_t^*)-\text{predictable and continuous } K\text{-a.e.}\},\\
S&=S_1\times S_2.
\end{aligned}$$

Processes in S_1 are identified if they agree at all times except perhaps for a $\mathbb{P}$-null set and processes in S_2 are identified if they agree K-a.e. If $T\in\mathcal{T}_b$, $\theta>0$ and $Z_1,Z_2\in S_2$, let

$$(V.4.2)\qquad d_{T,\theta}(Z_1,Z_2)=\hat{\mathbb{P}}_T\Big(\int_0^T(\sup_{u\le s}\|Z_1(u)-Z_2(u)\|^2\wedge 1)e^{-\theta s}ds\Big)^{1/2}\le\theta^{-1/2},$$

where, as usual, $\hat{\mathbb{P}}_T$ is the normalized Campbell measure associated with K_T. If $\bar\theta=(\theta_N,N\ge1)$ is a positive sequence satisfying

$$(V.4.3)\qquad \sum_1^\infty N\theta_N^{-1/2}<\infty,$$

define metrics $d_i=d_i^{\bar\theta}$ on S_i and $d_0=d_0^{\bar\theta}$ on S by

$$\begin{aligned}
d_1(X_1,X_2)&=\sum_{N=1}^\infty\mathbb{P}\Big(\int_0^{T_N}\sup_{u\le s}d(X_1(u),X_2(u))^2e^{-\theta_N s}ds\Big)^{1/2},\\
d_2(Z_1,Z_2)&=\sum_{N=1}^\infty\sup_{T\le T_N,T\in\mathcal{T}_b}d_{T,\theta_N}(Z_1,Z_2)<\infty\quad(\text{by (V.4.2)}),\\
d_0((X_1,Z_1),(X_2,Z_2))&=d_1(X_1,X_2)+d_2(Z_1,Z_2).
\end{aligned}$$

Note that if $u < T_N$, then $d(X_1(u), X_2(u)) \le X_1(u)(1) + X_2(u)(1) \le 2N$ and so by (V.4.3),

$$d_1(X_1, X_2) \le \sum_N 2N\theta_N^{-1/2} < \infty.$$

Lemma V.4.3. (S_i, d_i), $i = 1, 2$ and (S, d_0) are complete metric spaces.
Proof. This is straightforward. We will only show d_2 is complete. Suppose $\{Z_n\}$ is d_2 Cauchy. Let $N \in \mathbb{N}$. Then

$$\lim_{m,n\to\infty} \sup_{T \le T_N, T \in \mathcal{T}_b} \mathbb{E}\Big(\int \int_0^T \sup_{u \le s} \|Z_n(u) - Z_m(u)\|^2 \wedge 1 ds K_T(dy)\Big) = 0.$$

An application of the Section Theorem implies that

$$\sup_{t \le T_N} \int \int_0^t (\sup_{u \le s} \|Z_n(u) - Z_m(u)\|^2 \wedge 1) ds K_t(dy) \xrightarrow{\mathbb{P}} 0 \text{ as } m, n \to \infty.$$

Since $Z_n - Z_m$ is continuous K-a.e., this implies

$$\sup_{t \le T_N} \int \sup_{s \le t} \|Z_n(s) - Z_m(s)\|^2 \wedge 1 K_t(dy) \xrightarrow{\mathbb{P}} 0 \text{ as } m, n \to \infty.$$

A standard argument now shows that there is a subsequence $\{n_k\}$ so that
$(V.4.4)$
$$Z_{n_k}(s, \omega, y) \text{ converges uniformly in } s \le t \text{ for } K_t - \text{a.e. } y \ \text{ for all } t \ge 0 \ \mathbb{P} - \text{a.s.}$$

Now define
$$Z(s, \omega, y) = \begin{cases} \lim_{k\to\infty} Z_{n_k}(s, \omega, y) & \text{if it exists} \\ \mathbf{0} \in \mathbb{R}^d & \text{otherwise.} \end{cases}$$

Then Z is $(\hat{\mathcal{F}}_t^*)$-predictable and continuous K-a.e. (by (V.4.4)). Dominated Convergence easily shows that $d_2(Z_{n_k}, Z) \to 0$ as $k \to \infty$. ∎

Proof of Theorem 4.1(a) and Theorem 4.2.
For $i = 1, 2$ define $\Phi^i = (\Phi_1^i, \Phi_2^i) : S \to S$ by

$$(V.4.5) \quad \Phi_2^i(X, Z)(t) \equiv \tilde{Z}_t^i(\omega, y) = Z_0^i(y_0) + \int_0^t \sigma(X_s, Z_s) dy(s) + \int_0^t b(X_s, Z_s) ds$$

and

$$(V.4.6) \qquad \Phi_1^i(X, Z)(t)(\cdot) \equiv \tilde{X}_t^i(\cdot) = \int 1(\tilde{Z}_t^i(\omega, y) \in \cdot) K_t^i(dy).$$

In (V.4.5) the stochastic integral is that of Section 3 relative to K. Note that this integral also defines a version of the K^1 stochastic integral. To see this, note it is trivial for simple integrands and then one can approximate as in the construction of the K-stochastic integral to see this equivalence persists for all integrands in $D(d, d)$ (defined with respect to K). Remark V.1.1(a) implies

$$(V.4.7) \qquad \int_0^t K_s(\|\sigma(X_s, Z_s)\|^2 + \|b(X_s, Z_s)\|) ds$$

$$\leq \int_0^t K_s(1)(C(X_s(1))^2 + C(X_s(1)))ds < \infty \quad \forall t > 0 \text{ a.s.}$$

Corollary V.3.6 therefore may be used to show that $\tilde{X}$ is a.s. continuous. Note also that $\tilde{X}_t(1) = K_t(1) \leq N$ for $t < T_N$ and it follows that Φ^i takes values in S. Clearly a fixed point of Φ^i would be a solution of $(SE)_{Z_0^i, K^i}$.

To avoid writing factors of $m(1)^{-1}$, assume that $m(1) = 1$. Let $T \in \mathcal{T}_b$ satisfy $T \leq T_N$. Let $\tilde{Z}^i$ and $\tilde{X}^i$ be as in (V.4.5) and (V.4.6) for some $(X^i, Z^i) \in S$. Doob's strong L^2 inequality and Cauchy-Schwarz imply

$$\begin{aligned}
&\hat{\mathbb{P}}_T\Big(\sup_{s\leq t\wedge T} \|\tilde{Z}^1(s) - \tilde{Z}^2(s)\|^2 \wedge 1\Big) \\
&\quad \leq c\hat{\mathbb{P}}_T\Big(\|Z_0^1 - Z_0^2\|^2 \wedge 1 + \int_0^{t\wedge T} \|\sigma(X_s^1, Z_s^1) - \sigma(X_s^2, Z_s^2)\|^2 \\
&\qquad\qquad + N\|b(X_s^1, Z_s^1) - b(X_s^2, Z_s^2)\|^2 ds\Big) \\
&\quad \leq c\int \|Z_0^1 - Z_0^2\|^2 \wedge 1 dm_2 \\
&\qquad + cL(N)^2 N\hat{\mathbb{P}}_T\Big(\int_0^{t\wedge T} (d(X_s^1, X_s^2) + \|Z_s^1 - Z_s^2\|)^2 \wedge C(N)^2 ds\Big),
\end{aligned} \tag{V.4.8}$$

where in the last line we used Proposition V.2.4 (a), (Lip) and Remark V.1.1 (a). Note that if $T > 0$ then $T_N > 0$ and so $K_T(1) \leq \sup_{t\leq T_N} K_t(1) \leq N$. If $T = 0$, $K_T(1)$ may be bigger than N but the integral in (V.4.8) is then zero. Therefore

$$\begin{aligned}
&\hat{\mathbb{P}}_T\Big(\sup_{s\leq t\wedge T} \|\tilde{Z}^1(s) - \tilde{Z}^2(s)\|^2 \wedge 1\Big) \\
&\quad \leq c_N\Big[\int \|Z_0^1 - Z_0^2\|^2 \wedge 1 dm_2 + \mathbb{P}\Big(\int_0^{t\wedge T} d(X_s^1, X_s^2)^2 \wedge 1 ds\Big) \\
&\qquad + \hat{\mathbb{P}}_T\Big(\int_0^{t\wedge T} \|Z_s^1 - Z_s^2\|^2 \wedge 1 ds\Big)\Big].
\end{aligned} \tag{V.4.9}$$

Multiply the above inequality by $e^{-\theta t}$ and integrate t over $\mathbb{R}_+$ to conclude that

$$\begin{aligned}
&\sup_{T\leq T_N, T\in\mathcal{T}_b} d_{T,\theta}(\tilde{Z}^1, \tilde{Z}^2)^2 \\
&\quad \leq \frac{c_N}{\theta}\Big[\int \|Z_0^1 - Z_0^2\|^2 \wedge 1 dm_2 + \mathbb{P}\Big(\int_0^{T_N} (d(X_s^1, X_s^2)^2 \wedge 1)^2 e^{-\theta s} ds\Big) \\
&\qquad + \sup_{T\leq T_N, T\in\mathcal{T}_b} \hat{\mathbb{P}}_T\Big(\int_0^T (\|Z_s^1 - Z_s^2\|^2 \wedge 1) e^{-\theta s} ds\Big)\Big].
\end{aligned}$$

Therefore

$$\begin{aligned}
\sup_{T\leq T_N, T\in\mathcal{T}_b} d_{T,\theta}(\tilde{Z}^1, \tilde{Z}^2) &\leq \sqrt{\frac{c_N}{\theta}}\Big[\Big(\int \|Z_0^1 - Z_0^2\|^2 \wedge 1 dm_2\Big)^{1/2} \\
&\quad + \mathbb{P}\Big(\int_0^{T_N} (d(X_s^1, X_s^2)^2 \wedge 1)^2 e^{-\theta s} ds\Big)^{1/2}
\end{aligned} \tag{V.4.10}$$

$$+ \sup_{T \le T_N, T \in \mathcal{T}_b} d_{T,\theta}(Z^1, Z^2)\Big].$$

Take $\theta = \theta_N$ in (V.4.10), assume

$$(V.4.11) \qquad \delta_0 = \sum_{N=1}^{\infty} \sqrt{\frac{c_N}{\theta_N}} < \infty,$$

and sum the resulting inequality over N to conclude that

$$(V.4.12) \quad d_2(\tilde{Z}^1, \tilde{Z}^2) \le \delta_0 \Big[\Big(\int \|Z_0^1 - Z_0^2\|^2 \wedge 1 dm\Big)^{1/2} + d_1(X^1, X^2) + d_2(Z^1, Z^2)\Big].$$

Consider next a bound for $d_1(\tilde{X}^1, \tilde{X}^2)$. Let $\phi \in \mathrm{Lip}_1$. Then

$$\begin{aligned} |\tilde{X}_u^1(\phi) - \tilde{X}_u^2(\phi)|^2 &\le 2\Big[\Big(\int \phi(\tilde{Z}_u^1) - \phi(\tilde{Z}_u^2) dK_u^1\Big)^2 + \Big(\int \phi(\tilde{Z}_u^2) d(K_u^1 - K_u^2)\Big)^2\Big] \\ &\le 2\Big[\Big(\int \|\tilde{Z}_u^1 - \tilde{Z}_u^2\| \wedge 2 dK_u^1\Big)^2 + (K_u^2(1) - K_u^1(1))^2\Big]. \end{aligned}$$

Therefore if $N(u) = \int \sup_{s \le u} \|\tilde{Z}_s^1 - \tilde{Z}_s^2\| \wedge 1 K_u^1(dy)$, then

$$(V.4.13) \qquad \sup_{u \le t \wedge T_N} d(\tilde{X}_u^1, \tilde{X}_u^2)^2 \le 8\Big[\sup_{u \le t \wedge T_N} N(u)^2 + \sup_{u \le t \wedge T_N} (K_u^2(1) - K_u^1(1))^2\Big].$$

Claim N is an $(\mathcal{F}_t)$-submartingale. Let $U \le V$ be $(\mathcal{F}_t)$-stopping times. Then Remark V.2.5 (c) implies

$$\mathbb{E}(N(U)) = \mathbb{E}\Big(\int \sup_{s \le U} \|\tilde{Z}_s^1 - \tilde{Z}_s^2\| \wedge 1\, dK_V\Big) \le \mathbb{E}(N(V)).$$

As $N(t) \le K_t(1) \in L^1$, the claim follows by a standard argument. On $\{T_N > 0\}$ we have $\sup_{t \le T_N} K_t(1) \le N$. Therefore (V.4.13), Cauchy-Schwarz and Doob's strong L^2 inequality imply

$$\begin{aligned} &\mathbb{P}(1(T_N > 0) \sup_{u \le t \wedge T_N} d(\tilde{X}_u^1, \tilde{X}_u^2)^2) \\ &\quad \le c\mathbb{P}([N(t \wedge T_N)^2 + (K_{t\wedge T_N}^2(1) - K_{t \wedge T_N}^1(1))^2] 1(T_N > 0)) \\ &\quad \le c\mathbb{P}\Big(N \int \sup_{s \le t \wedge T_N} \|\tilde{Z}_s^1 - \tilde{Z}_s^2\|^2 \wedge 1\, dK_{t \wedge T_N}^2 + N(K_{t \wedge T_N}^2(1) - K_{t\wedge T_N}^1(1))\Big) \\ &\quad = cN\Big[\hat{\mathbb{P}}_{T_N}\Big(\sup_{s \le t \wedge T_N} \|\tilde{Z}_s^1 - \tilde{Z}_s^2\|^2 \wedge 1\Big) + m_2(1) - m_1(1)\Big] \quad \text{(by Remark V.2.5 (c))} \\ &\quad \le c_N'\Big[\int \|Z_0^1 - Z_0^2\|^2 \wedge 1\, dm_2 + m_2(1) - m_1(1) + \mathbb{P}\Big(\int_0^{t \wedge T_N} d(X_s^1, X_s^2)^2 \wedge 1\, ds\Big) \\ &\qquad\qquad + \hat{\mathbb{P}}_{T_N}\Big(\int_0^{t \wedge T_N} \|Z_s^1 - Z_s^2\|^2 \wedge 1\, ds\Big)\Big], \end{aligned}$$

the last by (V.4.9). It follows that

$$\begin{aligned}
&\mathbb{P}\Big(\int_0^{T_N} \sup_{u\le t} d(\tilde X^1_u, \tilde X^2_u)^2 e^{-\theta_N t} dt\Big)\\
&\quad\le \int_0^\infty \mathbb{P}\Big(1(T_N>0) \sup_{u\le t\wedge T_N} d(\tilde X^1_u, \tilde X^2_u)^2\Big) e^{-\theta_N t} dt\\
&\quad\le \frac{c'_N}{\theta_N}\Big[\int \|Z^1_0 - Z^2_0\|^2 \wedge 1\, dm_2 + m_2(1) - m_1(1)\\
&\qquad + \mathbb{P}\Big(\int_0^{T_N} (d(X^1_s, X^2_s)^2 \wedge 1)\, e^{-\theta_N s}\, ds\Big) + \hat{\mathbb{P}}_{T_N}\Big(\int_0^{T_N} (\|Z^1_s - Z^2_s\|^2 \wedge 1)\, e^{-\theta_N s}\, ds\Big)\Big].
\end{aligned}$$

Therefore if

$$\delta'_0 = \sum_1^\infty \sqrt{\frac{c'_N}{\theta_N}} < \infty, \tag{V.4.14}$$

then

$$\begin{aligned}
d_1(\tilde X^1, \tilde X^2) \le \delta'_0\Big(\int \|Z^1_0 - Z^2_0\|^2 \wedge 1\, dm_2^{1/2} + (m_2(1) - m_1(1))^{1/2}\\
+ d_1(X^1, X^2) + d_2(Z^1, Z^2)\Big).
\end{aligned} \tag{V.4.15}$$

Now set $Z^1_0 = Z^2_0 = Z_0$ and $K^1 = K^2 = K$ and $\Phi^i = \Phi$. (V.4.12) and (V.4.15) imply

$$d_0(\Phi(X^1, Z^1), \Phi(X^2, Z^2)) \le (\delta_0 + \delta'_0) d_0((X^1, Z^1), (X^2, Z^2)).$$

Therefore if we choose $\{\theta_N\}$ so that (V.4.11) and (V.4.14) hold with $\delta_0, \delta'_0 \le 1/4$, then Φ is a contraction on the complete metric space (S, d_0). It therefore has a unique fixed point (X, Z) which is a solution of $(SE)_{Z_0,K}$. Conversely if (X, Z) is a solution of $(SE)_{Z_0,K}$ then $X_s(1) = K_s(1)$ and, as in (V.4.7) and the ensuing argument, we see that $(X, Z) \in S$. Therefore (X, Z) is a fixed point of Φ and so is pathwise unique. This completes the proof of Theorem V.4.1 (a).

For Theorem 4.2, let (X^i, Z^i) satisfy $(SE)_{Z^i_0,K^i}$, $i = 1, 2$. Then $(X^i, Z^i) \in S$ by the above and so $\Phi^i(X^i, Z^i) = (X^i, Z^i)$. Therefore (V.4.12) and (V.4.15) imply

$$\begin{aligned}
d_0((X^1, Z^1), (X^2, Z^2)) \le &(\delta_0 + \delta'_0)\Big[d_0((X^1, Z^1), (X^2, Z^2))\\
&+ \Big(\int \|Z^1_0 - Z^2_0\|^2 \wedge 1\, dm_2\Big)^{1/2} + (m_2(1) - m_1(1))^{1/2}\Big].
\end{aligned}$$

As $\delta_0 + \delta'_0 \le 1/2$ by our choice of θ_N, this gives

$$\begin{aligned}
d_1(X^1, X^2) &\le d_0((X^1, Z^1), (X^2, Z^2))\\
&\le \Big(\int \|Z^1_0 - Z^2_0\|^2 \wedge 1\, dm_2\Big)^{1/2} + (m_2(1) - m_1(1))^{1/2},
\end{aligned}$$

and hence

$$\mathbb{P}\Big(\int_0^{T_N} \sup_{u\le s} d(X^1(u), X^2(u))^2\, ds\Big)$$
$$\le e^{\theta_N N} d_1(X^1, X^2)^2 \le 2e^{\theta_N N}\Big(m_2(1) - m_1(1) + \int \|Z_0^1 - Z_0^2\|^2 \wedge 1\, dm_2\Big). \quad \blacksquare$$

Proof of Theorem 4.1(b). Let $(\tilde{X}, \tilde{Z})$ be the unique solution of $(SE)_{Z_0,H}$, where H is the canonical process on $\bar{\Omega}_H = (\Omega_H, \mathcal{F}_H, \bar{\mathcal{H}}_t, \mathbb{Q}_{0,\nu})$. Clearly $(\tilde{X}, \tilde{Z})$ depends only on Z_0 and ν, and satisfies the required predictability conditions. We must show that $(\tilde{X}(t, K(\omega)), \tilde{Z}(t, K(\omega), y))$ solves $(SE)_{Z_0,K}$ on $\bar{\Omega}$. Let $I_H(f, t, H, y)$, $f \in D_H$ denote the stochastic integral from Proposition V.3.2 on $\bar{\Omega}_H$ and $I(f, t, \omega, y)$, $f \in D$ continue to denote that with respect to K on $\bar{\Omega}$. We claim that if $f \in D_H$, then $f \circ K(t, \omega, y) \equiv f(t, K(\omega), y) \in D$ and

$$(V.4.16) \qquad I(f \circ K) = I_H(f) \circ K \quad K - \text{a.e.}$$

The first implication is clear because K has law $\mathbb{Q}_{0,\nu}$. (V.4.16) is immediate if f is a $(\hat{\bar{\mathcal{H}}}_t)$-predictable simple function. A simple approximation argument using Proposition V.3.2 (d) then gives (V.4.16) for all $f \in D_H$. It is now easy to replace H with $K(\omega)$ in $(SE)_{Z_0,H}$ and see that $(\tilde{X} \circ K, \tilde{Z} \circ K)$ solves $(SE)_{Z_0,K}$ on $\bar{\Omega}$. $\blacksquare$

For (c) we need:

Lemma V.4.4. Let $\mu_1, \mu_2 \in M_F(\mathbb{R}^d)$ satisfy $\mu_1(\mathbb{R}^d) \le \mu_2(\mathbb{R}^d)$. There is a measure $m \in M_F(\mathbb{R}^d)$, $B \in \mathcal{B}(\mathbb{R}^d)$, and Borel maps $g_i : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\begin{aligned} &(1)\ \mu_1(\cdot) = m(g_1^{-1}(\cdot) \cap B),\ \ \mu_2(\cdot) = m(g_2^{-1}(\cdot)) \\ &(2)\ \int \|g_1 - g_2\| \wedge 1\, dm \le 2d(\mu_1, \mu_2) \\ &(3)\ m(B^c) = \mu_2(1) - \mu_1(1) \le d(\mu_1, \mu_2). \end{aligned}$$

Proof. (3) is immediate from (1).

If $\mu_1(\mathbb{R}^d) = \mu_2(\mathbb{R}^d)$, (1) and (2) with $B = \mathbb{R}^d$ and no factor of 2 in (2) is a standard "marriage lemma" (see, e.g., Szulga (1982)). Although the usual formulation has m defined on the Borel sets of $(\mathbb{R}^d)^2$ and g_i the projection maps from $(\mathbb{R}^d)^2$ onto $\mathbb{R}^d$, the above results follows as m and g_i can be carried over to $\mathbb{R}^d$ through a measure isomorphism between this space and $(\mathbb{R}^d)^2$.

If $\mu_1(\mathbb{R}^d) < \mu_2(\mathbb{R}^d)$, let $\mu_1' = \mu_1 + (\mu_2(\mathbb{R}^d) - \mu_1(\mathbb{R}^d))\delta_{x_0}$, where x_0 is chosen so that $\mu_1(\{x_0\}) = 0$. By the above case there are m, g_i satisfying (1) and (2) with μ_1' in place of μ_1, $B = \mathbb{R}^d$, and no factor of 2 in (2). (1) follows easily with $B = g_1^{-1}(x_0)^c$. For (2) note that

$$\begin{aligned} \int \|g_1 - g_2\| \wedge 1\, dm &\le d(\mu_1', \mu_2) \quad \text{(by the above case)} \\ &\le d(\mu_1', \mu_1) + d(\mu_1, \mu_2) \\ &= \mu_2(1) - \mu_1(1) + d(\mu_1, \mu_2) \le 2d(\mu_1, \mu_2). \quad \blacksquare \end{aligned}$$

Proof of Theorem 4.1(c). Assume first that $K_0 = m$ is deterministic and let (X, Z) solve $(SE)_{Z_0,K}$. The law of X on Ω_X, $\mathbb{P}_{m,Z_0}$, depends only on (m, Z_0) by (b). The next step is to show that it in fact only depends on $X_0 = m(Z_0^{-1}(\cdot))$. Define $\tilde{Z}_0(y)(t) = y(t) - y(0) + Z_0(y(0))$ and $K'_t(\phi) = K_t(\phi \circ \tilde{Z}_0)$. Then K' is an $(\mathcal{F}_t)$-historical Brownian motion starting at $(0, X_0)$ by Exercise V.2.1 with $B = \mathbb{R}^d$. Let (X', Z') be the unique solution of $(SE)_{Z'_0,K'}$, where $Z'_0(y_0) = y_0$. Let I, I' denote the stochastic integrals on $D(d, d)$ and $D'(d, d)$, respectively with respect to K and K', respectively. If $f : \mathbb{R}_+ \times \hat{\Omega} \to \mathbb{R}^{d\times d}$, set $f \circ \tilde{Z}_0(t, \omega, y) = f(t, \omega, \tilde{Z}_0(y))$. We claim that if $f \in D'$, then $f \circ \tilde{Z}_0 \in D$ and

$$(V.4.17) \qquad I'(f) \circ \tilde{Z}_0 = I(f \circ \tilde{Z}_0) \quad K - \text{a.e.}$$

The first inclusion is trivial. To prove the equality note it is obvious if f is simple and then use Proposition V.3.2(d) to extend the equality to all f in D' by approximating by simple functions as in the construction of I. If $\hat{Z}_t = Z'_t \circ \tilde{Z}_0$, then $\hat{Z}$ is $(\hat{\mathcal{F}}_t^*)$-predictable, and

$$Z'_t = y_0 + I'(\sigma(X', Z'), t) + \int_0^t b(X'_s, Z'_s)ds \quad K' - \text{a.e.}$$

implies

$$\hat{Z}_t = \tilde{Z}_0(y)(0) + I'(\sigma(X', Z'), t) \circ \tilde{Z}_0 + \int_0^t b(X'_s, Z'_s) \circ \tilde{Z}_0 \, ds \quad K - \text{a.e.}$$

Now use (V.4.17) to get

$$\hat{Z}_t = Z_0(y_0) + I(\sigma(X', \hat{Z}), t) + \int_0^t b(X'_s, \hat{Z}_s)ds \quad K - \text{a.e.},$$

and also note that $\mathbb{P}$-a.s. for all $t \geq 0$,

$$X'_t(\cdot) = \int 1(Z'_t \in \cdot)K'_t(dy) = \int 1(\hat{Z}_t \in \cdot)K_t(dy).$$

Therefore $(X', \hat{Z})$ solves $(SE)_{Z_0,K}$ and so $X' = X$ $\mathbb{P}$-a.s. by (a). This implies they have the same law and so $\mathbb{P}_{m,Z_0} = \mathbb{P}_{X_0,\mathrm{id}} \equiv \mathbb{P}'_{X_0}$, thus proving the claim.

To show the continuity of $\mathbb{P}'_{X_0}$ in X_0 we will use the stability of solutions to (SE) with respect to the initial conditions (Theorem V.4.2). Let $X_0^i \in M_F(\mathbb{R}^d)$, $i = 1, 2$ and choose m, B, and $g_i \equiv Z_0^i$ as in Lemma V.4.4 with $\mu_i = X_0^i$. Let $K_t^2 = K_t$ be an $(\mathcal{F}_t)$-historical Brownian motion starting at $(0, m)$ and define

$$K_t^1(A) = K_t(A \cap \{y : y(0) \in B\}).$$

Then Exercise V.2.1 shows that $K^1 (\leq K^2)$ is an $(\mathcal{F}_t)$-historical Brownian motion starting at $(0, m(\cdot \cap B))$. If (X^i, Z^i) solves $(SE)_{Z_0^i,K^i}$, then $X^i(0) = X_0^i$, as the notation suggests, and so X_i has law $\mathbb{P}'_{X_0^i}$. Introduce the uniform metric $\rho_M(x^1, x^2) = \sup_t d(x_t^1, x_t^2) \wedge 1$ on Ω_X and let d_M denote the corresponding Vasershtein metric on $M_F(\Omega_X)$. This imposes a stronger (i.e., uniform) topology on Ω_X, and hence on $M_F(\Omega_X)$, but as our processes have compact support in time the strengthening is illusory. If T_N is as in Theorem V.4.2 and $\zeta=\inf\{t:K_t(1)=0\}$, then $X^i_t=0$ if $t\ge\zeta$ and so

$$\begin{aligned}
d_M(\mathbb{P}'_{X_0^1},\mathbb{P}'_{X_0^2}) &\le \sup_{\phi\in\mathrm{Lip}_1}\int|\phi(X^1)-\phi(X^2)|\,d\mathbb{P}\\
&\le \int\sup_t d(X^1_t,X^2_t)\wedge 1\,d\mathbb{P}\\
&\le \mathbb{P}(T_N\le\zeta)+\mathbb{P}\Big(1(T_N>\zeta)\sup_{t\le T_N}d(X^1_t,X^2_t)\wedge 1\Big)\\
&\le \mathbb{P}(T_N\le\zeta)+\mathbb{P}\Big(\int_{T_N}^{T_{N+1}}\sup_{t\le u}d(X^1_u,X^2_u)^2\wedge 1\,du\,1(T_N>\zeta)\Big)^{1/2},
\end{aligned}$$

the last because on $\{T_N>\zeta\}$, $T_N=N$ and $T_{N+1}=N+1$. Theorem V.4.2 implies that

$$\begin{aligned}
d_M(\mathbb{P}'_{X_0^1},\mathbb{P}'_{X_0^2}) &\le \mathbb{P}(T_N\le\zeta)+\sqrt{c_{N+1}\Big(\int\|Z_0^1-Z_0^2\|^2\wedge 1\,dm+m(1)-m_1(1)\Big)}\\
&\le \mathbb{P}(T_N\le\zeta)+\sqrt{c_{N+1}3d(X^1_0,X^2_0)}.
\end{aligned}$$

The first term approaches zero as $N\to\infty$ and so the uniform continuity of $X_0\to\mathbb{P}'_{X_0}$ with respect to the above metrics on $M_1(\Omega_X)$ and $M_F(\mathbb{R}^d)$ follows.

Returning now to the general setting of (a) in which K_0 and X_0 may be random, we claim that if $A\in\mathcal{F}_X$, then

$$\mathbb{P}(X\in A|\mathcal{F}_0)(\omega)=\mathbb{P}'_{X_0(\omega)}(A)\quad \mathbb{P}-\text{a.s.} \tag{V.4.18}$$

Take expectations of both sides to complete the proof of (c). For (V.4.18), use (b) and the $(\mathcal{F}_t)$-Markov property of K to see that $\mathbb{P}$-a.s.,

$$\mathbb{P}(X\in A|\mathcal{F}_0)(\omega)=\mathbb{P}(\tilde X(K)\in A|\mathcal{F}_0)(\omega)=\mathbb{Q}_{0,K_0(\omega)}(\tilde X\in A). \tag{V.4.19}$$

Recall that $(\tilde X,\tilde Z)$ is the solution of $(SE)_{Z_0,H}$ on $(\Omega_H,\mathcal{F}_H,\bar{\mathcal{H}}_t,\mathbb{Q}_{0,\nu})$. Let $\bar{\mathcal{H}}^{K_0}$ be the augmentation of $\mathcal{F}^H[0,t+]$ with $\mathbb{Q}_{0,K_0}$-null sets. Claim that

$$\text{For } \nu-\text{a.a. } K_0,\quad (\tilde X,\tilde Z) \text{ satisfies } (SE)_{Z_0,H} \text{ on } (\Omega_H,\mathcal{F}_H,\bar{\mathcal{H}}_t,\mathbb{Q}_{0,K_0}). \tag{V.4.20}$$

The only issue is, as usual, the interpretation of $I(\sigma(\tilde X,\tilde Z),t)$ under these various measures. Let $I_{K_0}(f)$, $f\in D_{K_0}$ be this integral under $\mathbb{Q}_{0,K_0}$ and $I(f)$, $f\in D$ be the integral under $\mathbb{Q}_{0,\nu}$. Starting with simple functions and bootstrapping up as usual we can show for ν-a.a. K_0,

$$\begin{aligned}
&f\in D \text{ implies } f\in D_{K_0} \text{ and in this case}\\
&I(f,t,\omega,y)=I_{K_0}(f,t,\omega,y)\ \forall t\le u\ \text{ for } K_u-\text{a.a. } y\ \forall u\ge 0\ \mathbb{Q}_{0,K_0}-\text{a.s.}
\end{aligned}$$

It is now a simple matter to prove (V.4.20).

Since $X_0(\omega)(\cdot) = K_0(\omega)(Z_0^{-1}(\cdot))$ a.s. by $(SE)_{Z_0,K}$, (V.4.20) and the result for deterministic initial conditions established above, imply that

$$\mathbb{Q}_{0,K_0(\omega)}(\tilde{X} \in \cdot) = \mathbb{P}'_{X_0(\omega)}(\cdot) \quad \text{a.s.}$$

Use this in (V.4.19) to derive (V.4.18) and so complete the proof of (c).

To establish the strong Markov property (d) we need some notation and a preliminary result.

Notation. If $s \geq 0$ let $\hat{\theta}_s(y)(t) = y(t+s) - y(s)$ and $\mathcal{F}_t^{(s)} = \mathcal{F}_{s+t}$. If $Z_s(\omega, y)$ is $\hat{\mathcal{F}}_s^*$-measurable and $\tilde{Z}_s(\omega, y)(t) = Z_s(\omega, y) + \hat{\theta}_s(y)(t)$, for each $\phi : \mathbb{R}_+ \times \hat{\Omega} \to E$, let

$$\phi^{(s)}(t, \omega, y) = \phi(t - s, \omega, \tilde{Z}_s(\omega, y))1(t \geq s),$$

suppressing dependence on Z_s. If $B \in \mathcal{F}$ has positive $\mathbb{P}$ measure, let $\mathbb{P}_B(\cdot) = \mathbb{P}(\cdot|B)$.

Lemma V.4.5. Let $s \geq 0$, $B \in \mathcal{F}_s$ have positive $\mathbb{P}$ measure and assume Z_s is as above. Then

$$K_t^{(s)}(A) \equiv \int 1_A(\tilde{Z}_s(\omega, y))K_{t+s}(dy), \quad t \geq 0$$

is an $(\mathcal{F}_t^{(s)})$-historical Brownian motion under $\mathbb{P}_B$.
Moreover $\mathbb{P}_B(K_0^{(s)}(1)) \leq m(1)/\mathbb{P}(B) < \infty$.

The reader who has done Exercise V.2.1 should have no trouble believing this. The proof of (d) now proceeds along familiar lines. The continuity of $X_0 \to \mathbb{P}'_{X_0}$ allows us to assume $T \equiv s$ is constant by a standard approximation of T by finite-valued stopping times. We then must show that

$(V.4.21)$ There is a process $\hat{Z}$ so that $(X_{s+\cdot}, \hat{Z})$ solves $(SE)_{\hat{Z}_0,K^{(s)}}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t^{(s)}, \mathbb{P}_B)$, where $\hat{Z}_0(y) = y_0$ and $B \in \mathcal{F}_s$ satisfies $\mathbb{P}(B) > 0$.

If B is as above and $A \in \mathcal{F}_X$ then (V.4.21) and (V.4.1) show that

$$\mathbb{P}_B(X_{s+\cdot} \in A) = \int \mathbb{P}'_{X_s(\omega)}(A)d\mathbb{P}_B,$$

and this implies

$$\mathbb{P}(X_{s+\cdot} \in A|\mathcal{F}_s)(\omega) = \mathbb{P}'_{X_s(\omega)}(A) \text{ a.s.},$$

as required.

The proofs of Lemma V.4.5 and (V.4.21) are somewhat tedious and are presented in the Appendix at the end of the Chapter for completeness. It should be clear from our discussion of the martingale problem for X in Section 5 below that this is not the best way to proceed for this particular result.

Now return to the martingale problem for X. Recall that $a(\mu, x) = \sigma(\mu, x)\sigma(\mu, x)^*$ and

$$A_\mu\phi(x) = \sum_{i=1}^d \sum_{j=1}^d a_{ij}(\mu, x)\phi_{ij}(x) + \sum_{i=1}^d b_i(\mu, x)\phi_i(x), \quad \text{for } \phi \in C_b^2(\mathbb{R}^d).$$

The following exercise is a simple application of Itô's Lemma (Theorem V.3.5) and is highly recommended.

Exercise V.4.1. If (X, Z) satisfies $(SE)_{Z_0,K}$ show that

$$(MP)^{a,b}_{X_0} \quad \text{For all } \phi \in C_b^2(\mathbb{R}^d), \quad M_t^X(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(A_{X_s}\phi)ds$$

$$\text{is a continuous } (\mathcal{F}_t)\text{-martingale such that } \langle M^X(\phi)\rangle_t = \int_0^t X_s(\phi^2)ds.$$

We will comment on the uniqueness of solutions to $(MP)^{a,b}_{X_0}$ in the next section. Uniqueness in (SE) alone is enough to show weak convergence of the branching particle systems from Section V.1 to the solution X of (SE). Of course (Lip) continues to be in force.

Theorem V.4.6. Let $m \in M_F(\mathbb{R}^d)$ and let X^N be the solution of $(SE)_N$ constructed in Section V.1. If $X_0 = m(Z_0 \in \cdot)$, then

$$\mathbb{P}(X^N \in \cdot) \overset{w}{\Rightarrow} \mathbb{P}'_{X_0} \text{ on } D(\mathbb{R}_+, M_F(\mathbb{R}^d)) \text{ as } N \to \infty.$$

Sketch of Proof. Tightness and the fact that all limit points are supported on the space of continuous paths may be proved as in Section II.4. One way to prove that all limit points coincide is to take limits in $(SE)_N$ and show that all limit points do arise as solutions of $(SE)_{Z_0,K}$ for some K. More general results are proved by Lopez (1996) using the historical martingale problem (HMP) discussed in the next section.

5. Martingale Problems and Interactions

Our goal in this Section is to survey some of the recent developments and ongoing research in the martingale problem formulation of measure-valued diffusions. In order to use the uniqueness of solutions to $(SE)_{Z_0,K}$ (both pathwise and in law) to show that the associated martingale problem $(MP)^{a,b}_{X_0}$ is well-posed, we would have to show that it is possible to realize any solution X of $(MP)^{a,b}_{X_0}$ as part of a solution (X, Z) to $(SE)_{Z_0,K}$ for some historical Brownian motion K. In general it is not possible to recover K from X (see Barlow and Perkins (1994) for this result for ordinary super-Brownian motion) and so this appears to require a non-trivial enlargement of our space.

Donnelly and Kurtz (1999) were able to resolve the analogous problem in the setting of their exchangeable particle representations through an elegant application of a general result of Kurtz (1998) on filtered martingale problems.

Theorem V.5.1 (Donnelly-Kurtz (1999), Kurtz (1998)). If $X_0 \in M_F(\mathbb{R}^d)$, (σ, b) satisfies (Lip), and $a = \sigma\sigma^*$, then $(MP)^{a,b}_{X_0}$ is well-posed.

Discussion. Although I had originally planned to present these ideas in detail (the treatment in Section 6.5 of Donnelly and Kurtz (1999) is a bit terse), I will have to settle for a few (even more terse) remarks here.

First, the ideas of Section V.4 are readily adapted to the exchangeable particle representation of Donnelly and Kurtz. Indeed it is somewhat simpler as (SE) is replaced by a countable system of (somewhat unappealing) stochastic differential equations with jumps. This pathwise uniqueness leads to the uniqueness of the martingale problem for the generator $\mathcal{A}$ of their exchangeable infinite particle system, $(X_k(t))$, and $X_t(1)$, the total population size. The underlying population is $X_t \in M_F(\mathbb{R}^d)$ where for each fixed t, $\frac{X_t(\cdot)}{X_t(1)}$ is the deFinetti measure of $(X_k(t))$ and so

$$(V.5.1) \qquad X_t(\phi) = X_t(1) \lim_{N\to\infty} (N)^{-1} \sum_{k=1}^{N} \phi(X_k(t)) \quad \text{a.s.}$$

Of course X will satisfy $(MP)^{a,b}_{X_0}$ and these particular solutions will be unique as the richer structure from which they are defined is unique. On the other hand, given an arbitrary solution X of $(MP)^{a,b}_{X_0}$, one can introduce

$$\nu_t(\cdot) = \mathbb{E}\Big(\prod_1^\infty \frac{X_t(\cdot)}{X_t(1)} 1(X_t(1) \in \cdot)\Big).$$

This would be the one-dimensional marginals of $((X_k(t)), X_t(1))$ if such an exchangeable system existed. Some stochastic calculus shows that ν_t satisfies the forward equation associated with $\mathcal{A}$:

$$\nu_t(\phi) = \nu_0(\phi) + \int_0^t \nu_s(\mathcal{A}\phi)\, ds, \quad \phi \in \mathcal{D}(\mathcal{A}).$$

The key step is then a result of Kurtz (1998) (Theorem 2.7), earlier versions of which go back at least to Echevaria (1982). It gives conditions on $\mathcal{A}$, satisfied in our setting, under which any solution of the above forward equation are the one-dimensional marginals of the solution to the martingale problem associated with $\mathcal{A}$. In our setting this result produces the required $((X_k(t)), X_t(1))$ from which X can be recovered by (V.5.1). Here one may notice one of many simplifications we have made along the way–to obtain (V.5.1) from the martingale problem for $\mathcal{A}$ we need to introduce some side conditions to guarantee the fixed time exchangeability of the particle system. Hence one needs to work with a "restricted" martingale problem and a "restricted" forward equation in the above. This shows that every solution to $(MP)^{a,b}_{X_0}$ arises from such an exchangeable particle system and in particular is unique in law by the first step described above.

The methods of the previous section also extend easily to the historical processes underlying the solutions obtained there. Let d_H be the Vasershtein metric on $M_F(C)$ associated with the metric $\sup_t \|y_t - y'_t\| \wedge 1$ on C. Let $\mathcal{F}^H_t = \mathcal{F}^H[0, t+]$ and assume

$$\begin{aligned} &\hat\sigma : \mathbb{R}_+ \times \hat\Omega_H \to \mathbb{R}^{d\times d},\ \hat b : \mathbb{R}_+ \times \hat\Omega_H \to \mathbb{R}^d \text{ are } (\widehat{\mathcal{F}^H_t})\text{-predictable} \\ &\text{and for some nondecreasing function } L \text{ satisfy} \\ \text{(HLip)} \quad &\|\hat\sigma(t,J,y)\| + \|\hat b(t,J,y)\| \le L(t \vee \sup_{s\le t} J_s(1)) \text{ and} \\ &\|\hat\sigma(t,J,y) - \hat\sigma(t,J',y')\| + \|\hat b(t,J,y) - \hat b(t,J',y')\| \\ &\quad \le L(t \vee \sup_{s\le t} J_s(1) \vee \sup_{s\le t} J'_s(1)) \Big[\sup_{s\le t} d_H(J_s, J'_s) + \sup_{s\le t} \|y_s - y'_s\|\Big]. \end{aligned}$$

The historical version of (SE) is:

$$(HSE)_{Z_0,K}\,(a)\ Z_t(\omega,y)=Z_0(y)+\int_0^t\hat\sigma(s,J,y)dy(s)+\int_0^t\hat b(s,J,Z)ds\quad K-\text{a.e.}$$
$$(b)\ J_t(\omega)(\cdot)=\int 1(Z(\omega,y)^t\in\cdot)K_t(dy)\quad\forall t\ge 0\ \mathbb{P}-\text{a.s.}$$

As in Section V.4, a fixed point argument shows that solutions to $(HSE)_{Z_0,K}$ exist and are pathwise unique (Theorem 4.10 in Perkins (1995)). Recall the class D_{fd} of finite dimensional cylinder functions from Example V.2.8. If $\hat a_{ij}=\hat\sigma\hat\sigma^*_{ij}$, the corresponding generator is

$$A_J\phi(t,y)=\hat b(t,J,y)\cdot\nabla\phi(t,y)+\frac12\sum_{i=1}^d\sum_{j=1}^d\hat a_{ij}(t,J,y)\phi_{ij}(t,y),\quad \phi\in D_{fd}.$$

Assume for simplicity that $K_0=m$ is deterministic and hence so is $J_0=m(Z_0\in\cdot)$. It is again a simple exercise (cf. Exercise V.4.1) to show that the solution J of $(HSE)_{Z_0,K}$ satisfies

$$\forall\phi\in D_{fd}\quad J_t(\phi)=J_0(\phi)+\int_0^t\int A_J\phi(s,y)J_s(dy)ds+M_t(\phi),$$

$(HMP)^{\hat a,\hat b}_{J_0}$ where $M_t(\phi)$ is a continuous $(\mathcal{F}_t)$-martingale such that

$$\langle M(\phi)\rangle_t=\int_0^t J_s(\phi_s^2)ds.$$

The situation in (HSE) is now symmetric in that a historical process J is constructed from a given historical Brownian motion K. If $\hat a(t,J,y)$ is positive definite it will be possible to reconstruct K from J so that (HSE) holds (the main steps are described below) and so we have a means of showing that any solution of (HMP) does satisfy (HSE) and hence can derive:

Theorem V.5.2 (Perkins (1995)). Assume (HLip), $J_0\in M_F^0(C)$, and $\hat a=\hat\sigma\hat\sigma^*$ satisfies

$$\langle\hat a(t,J,y)v,v\rangle>0\ \forall v\in\mathbb{R}^d-\{0\}\text{ for }J_t-\text{a.a. }y\ \forall t\ge 0.$$

Then $(HMP)^{\hat a,\hat b}_{J_0}$ is well-posed.

One can use the change of measure technique in Section IV.1 to easily obtain the same conclusion for $(HMP)^{\hat a,\hat b,\hat g}_{J_0}$ in which $A_J\phi(t,y)$ is replaced with $A_J\phi(t,y)+\hat g(t,y)\phi(t,y)$, where $\hat g:\mathbb{R}_+\times\hat\Omega_H\to\mathbb{R}$ is bounded and $(\hat{\mathcal{F}}^H_t)$-predictable. As J is intrinsically time-inhomogeneous, one should work with general starting times $\tau\ge 0$ and specify $J_{t\wedge\tau}=J^0\in\{H\in\Omega_H: H_{\cdot\wedge\tau}=H\}\equiv\Omega_H^\tau$. The resulting historical martingale problem $(HMP)^{\hat a,\hat b,\hat g}_{\tau,J^0}$ is again well-posed and if $(\hat a,\hat b,\hat g)(t,J,y)=(\tilde a,\tilde b,\tilde g)(t,J_t,y)$, the solution will be a time-inhomogeneous $(\mathcal{F}_t)$-strong Markov process.

To prove the above results one needs to start with a solution of $(HMP)^{\hat a,\hat b}_{J^0}$, say, and develop the stochastic integration results in Section V.3 with respect to J. This general construction is carried out in Section 2 of Perkins (1995) and more general stochastic integrals for "historical semimartingales" with jumps may be found in Evans and Perkins (1998), although a general construction has not been carried out to date. Under the above non-degeneracy condition on $\hat{a}$, this then allows one define a historical Brownian motion, K, from J so that $(HSE)_{Z_0,K}$ holds, just as one can define the Brownian motion, B, from the solution, X, of $dX = \sigma(X)dB + b(X)dt$.

Consider next the problem of interactive branching rates. If $\gamma : M_F(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}_+$, then the extension of $(MP)^{a,b}_{X_0}$ which incorporates this state-dependent branching rate is

$$(MP)^{a,b,\gamma}_{X_0} \quad \forall\phi \in C^2_b(\mathbb{R}^d) \quad M^X_t(\phi) = X_t(\phi) - X_0(\phi) - \int_0^t X_s(A_{X_s}\phi)ds \text{ is a}$$
$$\text{continuous } (\mathcal{F}_t)\text{-martingale such that } \langle M^X(\phi)\rangle_t = \int_0^t X_s(\gamma(X_s)\phi^2)ds.$$

In general, uniqueness in law of X remains open. In the context of Fleming-Viot processes Dawson and March (1995) were able to use a dual process for the moments to resolve the analogous problem in which the sampling rates $\gamma(\mu, x, y)$ of types x and y may depend on the population μ in a smooth manner. Their result is a perturbation theorem analogous to that of Stroock and Varadhan (1979) for finite-dimensional diffusions but the rigidity of the norms does not allow one to carry out the localization step and so this very nice approach (so far) only establishes uniqueness for sampling rates which are close enough to a constant rate.

For our measure-valued branching setting, particular cases of state-dependent branching rates have been treated by special duality arguments (recent examples include Mytnik (1998) and Dawson, Etheridge, Fleischmann, Mytnik, Perkins, Xiong (2000a, 2000b)). If we replace $\mathbb{R}^d$ with the finite set $\{1, \ldots d\}$ and the generator A_μ with the state dependent Q-matrix $(q_{ij}(x))$, the solutions to the above martingale problem will be solutions to the stochastic differential equation

$$(V.5.2) \qquad dX^j_t = \sqrt{2\gamma_j(X_t)X^j_t}\,dB_j(t) + \sum_{i=1}^d X^i_t q_{ij}(X_t)dt.$$

Some progress on the uniqueness of solutions to this degenerate sde has recently been made by Athreya, Barlow, Bass, and Perkins (2000).

If $\hat{\gamma} : \mathbb{R}_+ \times \hat{\Omega}_H \to (0, \infty)$ is $(\hat{\mathcal{F}}^H_t)$-predictable, then conditions on $\hat{\gamma}$ are given in Perkins (1995) under which $(HMP)^{\hat{\gamma},\hat{a},\hat{b},\hat{g}}_{\tau,J_0}$ is well-posed. In this martingale problem we have of course

$$\langle M(\phi)\rangle_t = \int_\tau^t \int \hat{\gamma}(s, J, y)\phi(s, y)^2 J_s(dy)ds, \quad \phi \in D_{fd}.$$

Although the precise condition is complicated (see p. 48-49 in Perkins (1995)), it basically implies that $\hat{\gamma}$ should be represented by a (possibly stochastic) time integral. It is satisfied in the following examples.

Example V.5.3. (a) $\hat{\gamma}(t, J, y) = \gamma(t, y(t))$ for $\gamma \in C^{1,2}_b(\mathbb{R}_+ \times \mathbb{R}^d)$, γ bounded away from zero. In this case there is no interaction but branching rates may depend on space-time location something which our strong equation approach does not directly allow.

(b) (Adler's branching goats). $\hat{\gamma}(t,J,y)=\exp\Big\{-\int_0^t\int p_\varepsilon(y'_s-y_t)J_s(dy')e^{-\alpha(t-s)}ds\Big\}$. The branching rate at y_t is reduced if our goat-like particles have grazed near y_t in the recent past.

(c) (General time averaging). $\hat{\gamma}(t,J,y)=\int_{t-\varepsilon}^t f_\varepsilon(t-s)\gamma(s^+,J,y)ds$, where $f_\varepsilon:\mathbb{R}_+\to\mathbb{R}_+$ is C^1, $\mathrm{supp}(f_\varepsilon)\subset[0,\varepsilon)$, $\int_0^\varepsilon f_\varepsilon(s)ds=1$, and γ satisfies its analogue of (HLip).

Given the difficulties already present in the finite-dimensional case (V.5.2), resolving the uniqueness of solutions to $(MP)^{\gamma,a,b}_{X_0}$ or $(HMP)^{\hat{\gamma},\hat{a},\hat{b}}_{\tau,J_0}$ would appear to be an interesting open problem (see Metivier (1987)).

V.6. Appendix: Proofs of Lemma V.4.5 and (V.4.21)

We start by proving (IV.4.21), assuming the validity of Lemma V.4.5, and then address the proof of the latter.

Assume Z_s, B, $\mathbb{P}_B$, and $K^{(s)}$ are as in Lemma V.4.5. Let $I_s(f)$, $f\in D_s(d,d)\equiv D_s$ refer to the stochastic integral with respect $K^{(s)}$ on $(\Omega,\mathcal{F},\mathcal{F}_t^{(s)},\mathbb{P})$. There is some possible confusion here because of the other probabilities $\mathbb{P}_B$. Note, however, that if $I_{s,B}(f)$, $f\in D_{s,B}$ denotes the integral under $\mathbb{P}_B$, then

$$(V.6.1)\qquad D_s\subset D_{s,B}\text{ and for }f\in D_s\text{ we may take }I_{s,B}(f)=I_s(f).$$

Let $I(f)$, $f\in D$ continue to denote the stochastic integral with respect to K. The expression "$K^{(s)}$-a.e." will always mean with respect to $\mathbb{P}$ (not $\mathbb{P}_B$). With these clarifications, and the notation $\phi^{(s)}$ introduced prior to Lemma V.4.5, we have:

Lemma V.6.1. (a) If $\psi:\mathbb{R}_+\times\hat{\Omega}\to\mathbb{R}^d$ is $\mathcal{B}\times\hat{\mathcal{F}}^*$-measurable, then

$$\psi^{(s)}=0\quad K-\text{a.e. iff }\psi=0\quad K^{(s)}-\text{a.e.}$$

(b) If $f\in D_s$, then $f^{(s)}\in D$ and $I_s^{(s)}(f)=I(f^{(s)})\quad K-\text{a.e.}$

Proof. (a) is a simple exercise in using the definitions.

(b) The same is true for the first implication in (b). To check the equality, first let

$$f(t,\omega,y)=\sum_1^n f_i(\omega,y)1_{(t_{i-1},t_i]}(t)+f_0(\omega,y)1_{\{0\}}(t),$$
$$\text{where }f_i\in b\hat{\mathcal{F}}^*_{t_i+s},\ 0=t_0<t_1\ldots<t_n\le\infty.$$

Then

$$\begin{aligned}(V.6.2)\qquad I_s^{(s)}(f,t,\omega,y)&=I_s(f,t-s,\omega,\tilde{Z}_s(\omega,y))1(t\ge s)\\&=\sum_1^n f_i(\omega,\tilde{Z}_s(\omega,y))\cdot[\tilde{Z}_s(\omega,y)((t-s)\wedge t_i)\\&\qquad\qquad-\tilde{Z}_s(\omega,y)((t-s)\wedge t_{i-1})]1(t\ge s)\\&=\sum_1^n f_i(\omega,\tilde{Z}_s(\omega,y))\cdot[y(t\wedge(t_i+s))-y(t\wedge(t_{i-1}+s))].\end{aligned}$$

We also have

$$f^{(s)}(t,\omega,y)=\sum_1^n f_i(\omega,\tilde{Z}_s(\omega,y))1_{(s+t_{i-1},s+t_i]}(t)+f_0(\omega,\tilde{Z}_s(\omega,y))1_{\{s\}}(t),$$

and so the required result follows for such a simple f from the definition of $I(f^{(s)})$ and (V.6.2).

If $f\in D_s$, then as in (V.3.3) there are $(\widehat{\mathcal{F}_t^{(s)}})$-simple functions $\{f_k\}$ such that

$$\mathbb{P}\Big(K_k^{(s)}\Big(\int_0^k \|f_k(t)-f(t)\|^2dt>2^{-k}\Big)\Big)<2^{-k}.$$

We know $f^{(s)}, f_k^{(s)}\in D$ by the above and therefore for $k\geq s$,

$$\begin{aligned}
&\mathbb{P}\Big(K_k\Big(\int_0^k \|f_k^{(s)}(t)-f^{(s)}(t)\|^2dt>2^{-k}\Big)\Big)\\
&=\mathbb{P}\Big(K_{k-s}^{(s)}\Big(\int_0^{k-s} \|f_k(t)-f(t)\|^2dt>2^{-k}\Big)\Big)\\
&\leq\mathbb{P}\Big(K_k^{(s)}\Big(\int_0^k \|f_k(t)-f(t)\|^2dt>2^{-k}\Big)\Big)\quad\text{(use Remark V.2.5 (a))}\\
&<2^{-k}.
\end{aligned}$$

A double application of Proposition V.3.2 (d) now allows us to prove the required equality by letting $k\to\infty$ in the result for f_k. ∎

Proof of (V.4.21). Recall that (X,Z) is the solution of $(SE)_{Z_0,K}$. This gives us the Z_s which is used to define $K^{(s)}$ and $\phi^{(s)}$. Note that a solution to $(SE)_{\hat{Z}_0,K^{(s)}}$ with respect to $\mathbb{P}$ will also be a solution with respect to $\mathbb{P}_B$ (by (V.6.1)) and so we may assume that $B=\Omega$.

By Lemma V.4.5 and Theorem V.4.1 (a) there is a unique solution $(\hat{X},\hat{Z})$ to $(SE)_{\hat{Z}_0,K^{(s)}}$ on $(\Omega,\mathcal{F},\mathcal{F}_t^{(s)},\mathbb{P})$. Define

$$Z'_t(\omega,y)=\begin{cases}Z_t(\omega,y) & \text{if } t<s\\ \hat{Z}_t^{(s)}(\omega,y) & \text{if } t\geq s,\end{cases}$$

and

$$X'_t(\omega)=\begin{cases}X_t(\omega) & \text{if } t<s\\ \hat{X}_{t-s}(\omega) & \text{if } t\geq s\end{cases}\equiv\begin{cases}X_t(\omega) & \text{if } t<s\\ \hat{X}_t^{(s)}(\omega) & \text{if } t\geq s\end{cases}.$$

If $V(t)=\int_0^t b(\hat{X}_u,\hat{Z}_u)du$, then by $(SE)_{\hat{Z}_0,K^{(s)}}$,

$$\hat{Z}(t)=y_0+I_s(\sigma(\hat{X},\hat{Z}),t)+V(t)\quad K^{(s)}-\text{a.e.},$$

and so Lemma V.6.1 implies that K-a.e.,

$$\begin{aligned}\hat{Z}^{(s)}(t) &= Z_s + I_s^{(s)}(\sigma(\hat{X},\hat{Z}),t) + V^{(s)}(t)\\ &= Z_s + I(\sigma(\hat{X},\hat{Z})^{(s)},t) + \int_0^{t-s} b(\hat{X}_u(\omega),\hat{Z}_u(\omega,\tilde{Z}_s))du\, I(t\ge s)\\ &= Z_s + I(\sigma(\hat{X}^{(s)},\hat{Z}^{(s)})1(\cdot\ge s),t) + \int_s^t b(\hat{X}_u^{(s)},\hat{Z}_u^{(s)})du\, 1(t\ge s).\end{aligned}$$

It follows that K-a.e. for $t \ge s$,

$$\begin{aligned}Z'(t) &= Z_s + I(\sigma(X'_\cdot,Z'_\cdot)\,1(\cdot\ge s),t) + \int_s^t b(X'_u,Z'_u)du\\ &= Z_s + \int_s^t \sigma(X'_u,Z'_u)dy(u) + \int_s^t b(X'_u,Z'_u)du.\end{aligned}$$

Therefore we see that K-a.e.,

$$Z'(t) = Z_0 + \int_0^t \sigma(X'_u,Z'_u)du + \int_0^t b(X'_u,Z'_u)du,$$

first for $t \ge s$ by the above, and then for all $t \ge 0$ by the fact that (X,Z) solves $(SE)_{Z_0,K}$. Also $\mathbb{P}$-a.s. for all $t \ge s$

$$\begin{aligned}X'_t(\cdot) &= \int 1(\hat{Z}_{t-s}\in\cdot)K_{t-s}^{(s)}(dy)\\ &= \int 1(\hat{Z}_{t-s}(\omega,\tilde{Z}_s(\omega,y))\in\cdot)K_t(dy)\\ &= \int 1(Z'_t(\omega,y)\in\cdot)K_t(dy),\end{aligned}$$

and the above equality is trivial for $t < s$. We have shown that (X',Z') solves $(SE)_{Z_0,K}$ and so $X' = X$ a.s. This means that $X_{t+s} = \hat{X}_t$ for all $t \ge 0$ a.s. and (V.4.21) is proved. ∎

Proof of Lemma V.4.5. We will show that

$$(V.6.3) \qquad K_t^{(s)} \text{ satisfies } (HMP)_{0,K_0^{(s)}} \text{ on } (\Omega,\mathcal{F},\mathcal{F}_t^{(s)},\mathbb{P}).$$

It follows immediately that the same is true with respect to $\mathbb{P}_B$ and so the first result follows. The last inequality then follows trivially.

Assume first that $Z_s(\omega,y) = Z_s(y)$, $Z_s \in b\mathcal{C}_s$. If $\phi \in \mathcal{D}(\hat{A})$ and $n(t,y) = \phi(t,y) - \int_0^t \hat{A}\phi(r,y)dr$, then for $t \ge s$,

$$\phi^{(s)}(t,y) = n(t-s,\tilde{Z}_s(y)) + \int_s^t (\hat{A}\phi)^{(s)}(r,y)dr.$$

If $y_0 \in C^s$, $s \le u \le t$, and $G = \{y : y^s \in F_1,\ \hat\theta_s(y)^{u-s} \in F_2\}$ for some $F_i \in \mathcal{C}$, then $G \in \mathcal{C}_u$ and

$$\begin{aligned} P_{s,y_0}(n(t-s,\tilde Z_s)1_G) &= 1_{F_1}(y_0)P_{s,y_0}(n(t-s,Z_s(y_0)+\hat\theta_s(\cdot))1_{F_2}(\hat\theta_s(\cdot)^{u-s})) \\ &= 1_{F_1}(y_0)\int n(t-s,Z_s(y_0)+y)1_{F_2}(y^{u-s})dP^0(y) \\ &= 1_{F_1}(y_0)\int n(u-s,Z_s(y_0)+y)1_{F_2}(y^{u-s})dP^0(y) \\ &= P_{s,y_0}(n(u-s,\tilde Z_s)1_G), \end{aligned}$$

reversing the above steps in the last line. This shows that for any $m_s \in M_F^s(C)$, $\phi^{(s)} \in \mathcal{D}(\bar A_{s,m_s})$ and $\bar A_{s,m_s}(\phi^{(s)}) = (\hat A\phi)^{(s)}$. Now apply Proposition V.2.6 to the historical Brownian motion $\{K_t : t \ge s\}$ to conclude

$$\begin{aligned} K_t^{(s)}(\phi_t) &= \int \phi_{s+t}^{(s)} K_{s+t}(dy) \\ &= K_s(\phi_s^{(s)}) + \int_s^{s+t}\int \phi^{(s)}(r,y)dM(r,y) + \int_s^{s+t} K_r((\hat A\phi)_r^{(s)})dr. \qquad (V.6.4) \end{aligned}$$

That is,

$$\begin{aligned} \int \phi_t(\tilde Z_s(y))K_{t+s}(dy) = \int \phi_0(\tilde Z_s(y))K_s(dy) &+ \int_s^{s+t}\int \phi(r-s,\tilde Z_s(y))dM(r,y) \\ &+ \int_s^{s+t}\int \hat A\phi(r-s,\tilde Z_s(y))K_r(dy)dr\ \forall t \ge 0 \text{ a.s.} \qquad (V.6.5) \end{aligned}$$

If $Z_s(\omega,y) = \sum_{i=1}^n 1_{B_i}(\omega)Z^i(y)$ for $B_i \in \mathcal{F}_s$ and $Z^i \in b\mathcal{C}_s$, set $Z_s = Z^i$ in (V.6.5), multiply by $1_{B_i}(\omega)$ and sum over i to see that (V.6.5) remains valid if $Z_s(y)$ is replaced by $Z_s(\omega,y)$. Now pass to the pointwise closure of this class of $Z_s(\omega,y)$ and use the continuity of ϕ and $\hat A\phi$ to conclude that (V.6.5) remains valid if Z_s is $\hat{\mathcal{F}}_s^*$-measurable (Remark V.2.5 (c) allows us to pass from $\hat{\mathcal{F}}_s$ to $\hat{\mathcal{F}}_s^*$ on the right-hand side of (V.6.5) for all $t \ge 0$ simultaneously). Now reinterpret this general form of (V.6.5) as (V.6.4) and let $M^{(s)}(\phi)(t)$ denote the stochastic integral in (V.6.4). Then $M^{(s)}(\phi)(t)$ is an $(\mathcal{F}_t^{(s)})$-martingale and

$$\langle M^{(s)}(\phi)\rangle_t = \int_s^{s+t} K_r({\phi_r^{(s)}}^2)dr = \int_0^t K_r^{(s)}(\phi_r^2)dr.$$

The last term on the right-hand side of (V.6.5) equals $\int_0^t K_r^{(s)}((\hat A\phi)_r)\,dr$ and the first term equals $K_0^{(s)}(\phi_0)$. This proves (V.6.3) and so completes the proof. ∎

References

R. Adler and R. Tribe (1998). Uniqueness for a historical SDE with a singular interaction, J. Theor. Prob. 11, 515-533.

D. Aldous (1993). Tree-based models for random distribution of mass. J. Stat. Phys. 73, 625-641.

S. Athreya, M. T. Barlow, R. Bass and E. Perkins (2000). Degenerate stochastic differential equations and super-Markov chains. Preprint.

P. Baras and M. Pierre (1984). Singularités éliminables pur des equations semilinéaires. Ann. Inst. Fourier 34, 185-206.

M.T. Barlow, S.N. Evans and E. Perkins (1991). Collision local times and measure-valued diffusions. Can. J. Math. 43, 897-938.

M.T. Barlow and E. Perkins (1994). On the filtration of historical Brownian motion. Ann. Prob. 22, 1273-1294.

M. Bramson, J.T. Cox, and J.F. Le Gall (2001). Super-Browian limits of voter model clusters. To appear, Ann. Prob.

M. Bramson, R. Durrett and G. Swindle (1989). Statistical mechanics of crabgrass. Ann. Prob. 17, 444-481.

D.L. Burkholder (1973). Distribution function inequalities for martingales. Ann. Prob. 1, 19-42.

Z. Ciesielski and S.J. Taylor (1962). First passage times and sojourn times for Brownian motion and exact Hausdorff measure of the sample path. Trans. Amer. Math. Soc. 103, 434-450.

J.T. Cox, R. Durrett, E. Perkins (1999). Rescaled particle systems converging to super-Brownian motion. In: Perplexing Problems in Probability–Festschrift in Honor of Harry Kesten, pp. 269-284, Birkhäuser, Boston.

J.T. Cox, R. Durrett, E. Perkins (2000). Rescaled voter models converge to super-Brownian motion. Ann. Prob. 28, 185-234.

D.A. Dawson (1978). Geostochastic calculus. Can. J. Statistics 6, 143-168.

D.A. Dawson (1992). Infinitely divisible random measures and superprocesses. In: Stochastic Analysis and Related Topics, pp. 1-130, Birkhäuser, Boston.

D. A. Dawson (1993). Measure-valued Markov Processes, Ecole d'Eté de Probabilités de Saint Flour 1991, Lect. Notes in Math. 1541, Springer, Berlin.

D. A. Dawson, A. Etheridge, K. Fleischmann, L. Mytnik. E. Perkins, and J. Xiong (2000a). Mutually catalytic branching in the plane. Preprint.

D. A. Dawson, K. Fleischmann, L. Mytnik. E. Perkins, and J. Xiong (2000b). Mutually catalytic branching in the plane: uniqueness. Preprint.

D.A. Dawson, I. Iscoe, E. Perkins (1989). Super-Brownian motion: path properties and hitting probabilities. Probab. Th. Rel. Fields 83, 135-205.

D.A. Dawson and K.J. Hochberg (1979). The carrying dimension of a stochastic measure diffusion. Ann. Prob. 7, 693-703.

D.A. Dawson and P. March (1995). Resolvent estimates for Fleming-Viot operators and uniqueness of solutions to related martingale problems. J. Funct. Anal. 132, 417-472.

D.A. Dawson, E. Perkins (1991). Historical Processes. Mem. Amer. Math. Soc. 93 n. 454.

D.A. Dawson and V. Vinogradov (1994). Almost sure path properties of $(2, d, \beta)$-superprocesses. Stoch. Proc. Appl. 51, 221-258.

C. Dellacherie and P.A. Meyer (1978). Probabilities and Potential. North-Holland Math. Studies 29, North Holland, Amsterdam.

E. Derbez and G. Slade (1998). The scaling limit of lattice trees in high dimensions. Commun. Math. Phys. 193, 69-104.

J.S. Dhersin and J.F. Le Gall (1997). Wiener's test for super-Brownian motion and for the Brownian snake. Probab. Th. Rel. Fields 108, 103-129.

P. Donnelly and T.G. Kurtz (1999). Particle representations for measure-valued population models. Ann. Prob. 27, 166-205.

J. Dugundji (1966). Topology. Allyn and Bacon, Boston.

R. Durrett and E. Perkins (1999). Rescaled contact processes converge to super-Brownian motion for $d \geq 2$. Probab. Th. Rel. Fields 114, 309-399.

E.B. Dynkin (1991). A probabilistic approach to one class of nonlinear differential equations. Probab. Th. Rel. Fields 89, 89-115.

E.B. Dynkin (1993). Superprocesses and partial differential equations. Ann. Prob. 20, 1185-1262.

E.B. Dynkin (1994). An Introduction to Branching Measure-Valued Processes, CRM Monographs 6, Amer. Math. Soc., Providence.

E.B. Dynkin and S.E. Kuznetsov (1996). Superdiffusions and removable singulariities for quasilinear partial diferential equations. Comm. Pure. Appl. Math. 48,125-176.

E.B. Dynkin and S.E. Kuznetsov (1998). Trace on the boundary for solutions of nonlinear differential equations. Trans. Amer. Math. Soc. 350, 4499-4519.

P.E. Echevaria (1982). A criterion for invariant measures of Markov processes. Z. f. Wahrsch. verw. Gebiete 61, 1-16.

N. El Karoui and S. Roelly (1991). Propriétés de martingales, explosion et représentation de Lévy-Khintchine d'une classe de processus de branchement à valeurs mesures. Stoch. Process. Appl. 38, 239-266.

A. Etheridge (2000). An Introduction to Superprocesses, University Lecture Series vol.20, Amer. Math. Soc.

A. Etheridge and P. March (1991). A note on superprocesses. Probab. Th. Rel. Fields 89, 141-147.

S.N. Ethier and T.G. Kurtz (1986). Markov Processes: Characterization and Convergence. Wiley, N.Y.

S.N. Ethier, and T.G. Kurtz (1993). Fleming-Viot processes in population genetics. Siam. J. Cont. Opt. 31, 345-386.

S.N. Evans and E. Perkins (1991). Absolute continuity results for superprocesses. Trans. Amer. Math. Soc. 325, 661-681.

S.N. Evans and E. Perkins (1994). Measure-valued branching diffusions with singular interactions. Can. J. Math. 46, 120-168.

S.N. Evans and E. Perkins (1998). Collision local times, historical stochastic calculus and competing species. Elect. J. Prob. 3, paper 5.

W. Feller (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in Wahrscheinlichkeitstheoretischer Behandlung. Acta Biotheoretica 5, 11-40.

W. Feller (1951). Diffusion processes in genetics. Proc. Second Berkeley Symp. Math. Statist. Prob., Univ. of California Press Berkeley, pp. 227-246.

P.J. Fitzsimmons (1988). Construction and regularity of measure-valued Markov processes. Israel J. Math. 64, 337-36.

P.J. Fitzsimmons (1992). On the martingale problem for measure-valued branching processes. In: Seminar on Stochastic Processes 1991, pp. 39-51, Birkhäuser, Boston.

I.I. Gihman and A.V. Skorohod (1975). The Theory of Stochastic Processes II. Springer, New York.

T.E. Harris (1963). The Theory of Branching Processes. Springer, Berlin.

R. van der Hofstad, G. Slade (2000). Convergence of critical oriented percolation to super-Brownian motion above 4+1 dimensions. Preprint.

D.N. Hoover and H.J. Keisler (1984). Adapted probability distributions. Trans. Amer. Math. Soc. 286, 159-201.

D.N. Hoover and E. Perkins (1983). Nonstandard construction of the stochastic integral and applications to stochastic differential equations I, II, Trans. Amer. Math. Soc. 275, 1-36, 37-58.

N. Ikeda and S. Watanabe (1981). Stochastic Differential Equations and Diffusion Processes. North Holland, Amsterdam.

I. Iscoe (1986). A weighted occupation time for a class of measure-valued branching processes. Probab. Th. Rel. Fields 71, 85-116.

I. Iscoe (1988). On the supports of measure-valued critical branching Brownian motion. Ann. Prob. 16, 200-221.

J. Jacod (1977). A general theorem of representation for martingales. In: Proceedings of Symposia in Pure Math. 31, pp. 37-54, Amer. Math. Soc, Providence.

J. Jacod and A.N. Shiryaev (1987). Limit theorems for Stochastic Processes. Springer-Verlag, Berlin.

A. Jakubowski (1986). On the Skorohod topology. Ann. Inst. H. Poincaré B22, 263-285.

S. Kakutani (1944). On Brownian motion in n-space. Proc. Imp. Acad. Tokyo 20, 648-652.

O. Kallenberg (1986). Random Measures. Academic Press, New York.

F. Knight (1981). Essentials of Brownian Motion and Diffusion. Amer. Math. Soc., Providence.

N. Konno and T. Shiga (1988). Stochastic differential equations for some measure-valued diffusions. Probab. Th. Rel. Fields 79, 201-225.

N. Krylov (1997). On a result of Mueller and Perkins. Probab. Th. Rel. Fields 108, 543-557.

N. Krylov (1997b). On SPDE's and superdiffusions. Ann. Prob. 25, 1789-1809.

T.G. Kurtz (1975). Semigroups of conditioned shifts and approximation of Markov processes. Ann. Prob. 3, 618-642.

T.G. Kurtz (1998). Martingale problems for conditional distributions of Markov processes. Elect. J. Prob. 3, paper 9.

J.F. Le Gall (1993). A class of path-valued Markov processes and its applications to superprocesses. Probab. Th. Rel. Fields 95, 25-46.

J.F. Le Gall (1993b). The uniform random tree in a Brownian excursion. Probab. Th. Rel. Fields 96, 369-383.

J. F. Le Gall (1994). A lemma on super-Brownian motion with some applications. In: The Dynkin Festschrift pp. 237-251. Birkhäuser, Boston.

J. F. Le Gall (1998). The Hausdorff measure of the range of super-Brownian motion. In: Perplexing Problems in Probability–Festschrift in Honor of Harry Kesten, pp. 285-314, Birkhäuser, Boston.

J. F. Le Gall (1999). Spatial Branching Processes, Random Snakes and Partial Differential Equations. Lectures in Mathematics ETH Zurich, Birkhäuser Verlag, Basel.

J. F. Le Gall and E. Perkins (1995). The Hausdorff measure of the support of two-dimensional super-Brownian motion. Ann. Prob. 232, 1719-1747.

J. F. Le Gall, E. Perkins and S.J. Taylor (1995). The packing measure of the support of super-Brownian motion. Stoch. Process. Appl. 59, 1-20.

Z. Li and T. Shiga (1995). Measure-valued branching diffusions: immigrations, excursions and limit theorems, J. Math. Kyoto Univ. 35, 233-274.

M. Lopez (1996). Path properties and convergence of interacting superprocesses. Ph.D. thesis, UBC.

M. Metivier (1987). Weak convergence of measure-valued processes using Sobolev-imbedding techniques. In: Proc. Trento 1985 SPDE and Applications, Lect. Notes. Math. 1236, 172-183, Springer, Berlin.

C. Mueller and E. Perkins (1992). The compact support property for solutions of the heat equation with noise. Probab. Th. Rel. Fields 93, 325-358.

C. Mueller and R. Tribe (1994). Stochastic pde's arising from the long range contact and long range voter processes. Probab. Th. Rel. Fields 102, 519-546.

L. Mytnik (1998). Weak uniqueness for the heat equation with noise. Ann. Prob. 26, 968-984.

L. Mytnik (1999). Uniqueness for a competing species model. Can. J. Math. 51, 372-448.

A. Pazy (1983). Semigroups of Linear Operators and Applications to Partial Differential Applications. Springer-Verlag, Berlin.

E. Perkins (1988). A space-time property of a class of measure-valued branching diffusions. Trans. Amer. Math. Soc. 305, 743-795.

E. Perkins (1989). The Hausdorff measure of the closed supoprt of super-Brownian motion. Ann. Inst. H. Poincaré Stat. 25, 205-224.

E. Perkins (1990). Polar sets and multiple points for super-Brownian motion. Ann. Prob. 18, 453-491.

E. Perkins (1991). On the continuity of measure-valued processes. In: Seminar on Stochastic Processes 1990, pp. 261-268, Birkhäuser, Boston.

E. Perkins (1992). Measure-valued branching diffusions with spatial interactions. Probab. Th. Rel. Fields 94, 189-245.

E. Perkins (1992b). Conditional Dawson-Watanabe processes and Fleming-Viot processes. In: Seminar on Stochastic Processes 1991, pp.142-155, Birkhäuser, Boston.

E. Perkins (1994). The Strong Markov Property of the Support of Super-Brownain Motion. In: The Dynkin Festschrift Markov Processes and their Applications, pp.307-326, Birkhäuser, Boston.

E. Perkins (1995). On the martingale problem for interactive measure-valued diffusions. Mem. Amer. Math. Soc. 115 n. 549.

E. Perkins (1995b). Measure-valued branching diffusions and interactions. Proceedings of the International Congress of Mathematicians, Zurich, 1994, pp.1036-1046, Birkhäuser Verlag, Basel.

E. Perkins and S.J. Taylor (1998). The multifractal spectrum of super-Brownian motion. Ann. Inst. H. Poincaré Stat. 34, 97-138.

S.C. Port and C. J. Stone (1978). Brownian Motion and Classical Potential Theory. Academic, New York.

M. Reimers (1989). One-dimensional stochastic partial differential equations and the branching measure-diffusion. Probab. Th. Rel. Fields 81, 319-340.

M. Reimers (1989b). Hyperfinite methods applied to the critical branching diffusion. Probab. Th. Rel. Fields 81, 11-27.

D. Revuz and M. Yor (1991). Continuous Martingales and Brownian Motion. Springer, Berlin.

S. Roelly-Coppoletta (1986). A criterion of convergence of measure-valued processes: application to measure branching processes. Stochastics 17,43-65.

L.C.G. Rogers and D. Williams (1987). Diffusions, Markov Processes and Martingales Vol. 2. Wiley, Chichester.

C. A. Rogers and S. J. Taylor (1961). Functions continuous and singular with respect to a Hausdorff measure. Mathematika 8, 1-31.

H. Rost (1971). The stopping distributions of a Markov process. Invent. Math. 14, 1-16.

W. Rudin (1974). Real and Complex Analysis. McGraw-Hill, New York.

A. Schied (1996). Sample path large deviations for super-Brownian motion. Probab. Th. Rel. Fields 104, 319-347.

M. Sharpe (1988). General Theory of Markov Processes. Academic, Boston.

T. Shiga (1990). A stochastic equation based on a Poisson system for a class of measure-valued diffusion processes. J. Math. Kyoto. Univ. 30, 245-279.

D.W. Stroock and S.R.S. Varadhan (1979). Multidimensional Diffusion Processes, Springer-Verlag, Berlin.

S. Sugitani (1987). Some properties for the measure-valued diffusion process. J. Math. Soc. Japan 41, 437-462.

A. Szulga (1982). On minimal metrics in the space of random variables. Teor. Veroyatn. Primen 27, 401-405.

S.J. Taylor (1961). On the connections between generalized capacities and Hausdorff measures. Proc. Cambridge Philos. Soc. 57, 524-531.

S.J. Taylor (1964). The exact Hausdorff measure of the sample path for planar Brownian motion. Proc. Cam. Phil. Soc. 60, 253-258.

S.J. Taylor and N.A. Watson (1985). A Hausdorff measure classification of polar sets for the heat equation. Math. Proc. Cam. Phil. Soc. 47, 325-344.

R. Tribe (1991). The connected components of the closed support of super-Brownian motion. Probab. Th. Rel. Fields 84, 75-87.

R. Tribe (1992). The behavior of superprocesses near extinction. Ann. Prob. 20, 286-311.

R. Tribe (1994). A representation for super Brownian motion, Stoch. Proc. Appl. 51, 207-219.

J. Walsh (1986). An Introduction to Stochastic Partial Differetial Equations, Ecole d'Eté de Probabilités de Saint Flour 1984, Lect. Notes. in Math. 1180, Springer, Berlin.

S. Watanabe (1968). A limit theorem of branching processes and continuous state branching processes, J. Math. Kyoto U. 8, 141-167.

T. Yamada and S. Watanabe (1971). On the uniqueness of solutions of stochastic differential equations. J. Math. Kyoto U. 11, 155-167.

Index