

REGULARITY OF INVARIANT DENSITIES FOR 1D-SYSTEMS WITH RANDOM SWITCHING

YURI BAKHTIN, TOBIAS HURTH, AND JONATHAN C. MATTINGLY

ABSTRACT. This is a detailed analysis of invariant measures for one-dimensional dynamical systems with random switching. In particular, we prove smoothness of the invariant densities away from critical points and describe the asymptotics of the invariant densities at critical points.

Keywords: randomly switched ODEs, piecewise deterministic Markov processes, invariant densities

MSC numbers: 93E15, 93C30, 37A50, 60J25

1. INTRODUCTION

In this paper, we study the regularity theory for invariant densities of dynamical systems with random switching (switching systems, in short) with one-dimensional continuous component. Dynamical systems with random switching are also known as piecewise deterministic Markov processes.

We show that smoothness of the vector fields governing the dynamics translates into smoothness of invariant densities away from critical points of the vector fields (Theorem 1). For analytic vector fields, we derive the asymptotically dominant term of an invariant density as its argument approaches a critical point of the corresponding vector field (Theorem 2).

In the literature, regularity properties of invariant densities are often assumed in order to derive other features of the densities. For instance, it is shown in [FGRC09, Proposition 3.1] that if invariant densities are \mathcal{C}^1 on a set Ω , they satisfy the Fokker–Planck equations associated to the switching system in the interior of Ω . From this differential characterization, the authors deduce time-reversibility of stationary piecewise deterministic Markov processes and derive explicit formulas for the invariant densities of certain switching systems that they call exactly solvable. A result similar to [FGRC09, Proposition 3.1] can be found in [HKNT98, Theorem 1]. In this paper, we give sufficient conditions for continuity and differentiability of invariant densities that are stated in terms of the vector fields, and are easily verifiable. In particular, we show that if none of the vector fields vanish at a point $\xi \in \mathbb{R}$ and if all vector fields are \mathcal{C}^{n+1} in a neighborhood of ξ , then the invariant densities are \mathcal{C}^n at ξ . This statement is not surprising, but as far as we know, no rigorous proof of it has been given yet.

The second question addressed in this paper is how invariant densities behave at critical points of the vector fields. In the case of two vector fields on a bounded interval that point in opposite directions, [FGRC09, Proposition 3.12] gives an explicit formula for the invariant densities. From this formula, one obtains the

exact asymptotic behavior of the densities close to critical points. However, computing invariant densities explicitly is in general very difficult ([FGRC09, Section 3.3]). Finding necessary and sufficient conditions for boundedness of invariant densities is already challenging. In the one-dimensional case, invariant densities are bounded away from critical points (Lemma 5), but we expect to find switching systems with two-dimensional continuous component whose invariant densities become unbounded along curves that do not contain any critical points. If the continuous component is one-dimensional, [BHK⁺11, Theorem 1] provides sufficient conditions for boundedness of an invariant density close to a critical point of its associated vector field. For vector fields that behave linearly close to a critical point, we give necessary conditions and sufficient conditions for boundedness in terms of the vector fields and the switching rates (Corollary 1). These conditions recover part of the results in [BHK⁺11]. For analytic vector fields, we also compute the asymptotically dominant term of an invariant density as its argument approaches a critical point of the associated vector field (Theorem 2). Even if the vector fields are not analytic, we can derive the asymptotically dominant term in certain cases (Theorem 3).

The basic tools in our investigation are two integral equations satisfied by invariant densities. These equations are closely related to the Fokker–Planck equations (see Appendix A), but do not require differentiability of the densities. When deriving the asymptotically dominant terms in the case of analytic vector fields, we use the theory of regular singular points for systems of linear ordinary differential equations. We follow [Tay11, Section 3.11].

The paper is organized as follows. In Section 2, we describe the class of switching systems we consider and introduce some notation and terminology needed in the rest of the paper. Our main results are stated in Section 3. In Section 4, we formulate the integral equations mentioned earlier. One of these equations plays an important role in the proof of Theorem 1 that can be found in Section 5. In Section 6, we describe the support of invariant measures for switching systems with one-dimensional continuous component. We exploit this description of the support in Section 7, which is devoted to proving Theorems 2 and 3. Section 8 contains the proofs of the integral equations from Section 4. The appendix contains a remark on how the integral equations relate to the Fokker–Planck equations for the invariant densities.

Acknowledgements. JCM acknowledges the support of the National Science Foundation(DMS-0854879) and the Simons Foundation.

2. DEFINITIONS AND NOTATION

We consider a similar setup as in [BH12], but we restrict our analysis to switching systems with one-dimensional continuous component. Let D be a finite collection of continuously differentiable and forward-complete vector fields on \mathbb{R} . We denote these vector fields by u_i , where i is some element of the index set $S = \{1, \dots, n\}$.

Our assumptions on the vector fields imply that the initial-value problem

$$\begin{aligned}\dot{x}(t) &= u_i(x(t)), \\ x(0) &= \xi\end{aligned}$$

induced by $u_i \in D$ and by $\xi \in \mathbb{R}$ has a unique solution defined for all $t \geq 0$. We define a stochastic process $X = (X_t)_{t \geq 0}$ on \mathbb{R} as follows. Given an initial point $\xi \in \mathbb{R}$

and an initial driving vector field $u_i \in D$, the process X follows the solution trajectory of the corresponding initial-value problem for an exponentially distributed random time, with intensity parameter $\lambda_i > 0$. Then, a new driving vector field u_j is selected at random from $D \setminus \{u_i\}$, and X follows the corresponding trajectory for another exponentially distributed random time, with intensity parameter $\lambda_j > 0$. Using exponential times is required to make the resulting process Markovian. This construction is iterated to obtain a continuous trajectory on \mathbb{R} that is defined for any time $t \geq 0$ and driven by exactly one of the vector fields in D between any two switches. We call the intensity parameters $(\lambda_i)_{i \in S}$ switching rates. For $j \neq i$, let $\lambda_{i,j}$ be the rate of switching from u_i to u_j . Then,

$$\lambda_i = \sum_{j \neq i} \lambda_{i,j}.$$

As in [BH12], we assume that the exponential times between any two switches are independent conditioned on the sequence of driving vector fields, that the switching rate λ_i depends only on the current state i (in particular not on the value of X at the given time), and that for any two states $i, j \in S$ there is a positive probability of switching from i to j . In many papers on dynamical systems with random switching, the switching rates are allowed to depend on the location of the process X , and it is only required that the transition mechanism on S be irreducible (see for instance [FGRC09], [BLBMZ12a] and [CH13]). We hope to simplify our exposition by not studying more general classes of switching systems.

While X alone is not Markov, we can build a Markov process by adjoining a second stochastic process $A = (A_t)_{t \geq 0}$ that captures the current driving vector field at any given time. The product space $\mathbb{R} \times S$ is then the state space of the 2-component Markov process (X, A) with Markov semigroup $(\mathbf{P}^t)_{t \geq 0}$. We denote the associated distributions on paths starting at points (ξ, i) by $\mathbf{P}_{\xi,i}$, and the corresponding transition probability measures by $\mathbf{P}_{\xi,i}^t, t \geq 0$. We call X the continuous and A the discrete component of (X, A) .

Recall that a probability measure μ on $\mathbb{R} \times S$ is called an invariant measure of $(\mathbf{P}^t)_{t \geq 0}$ if

$$(1) \quad \mu(E \times \{i\}) = \mu \mathbf{P}^t(E \times \{i\}) := \sum_{j \in S} \int_{\mathbb{R}} \mathbf{P}_{\xi,j}^t(E \times \{i\}) \mu(d\xi \times \{j\})$$

holds for any Borel set $E \subset \mathbb{R}$, for any $i \in S$ and for any $t \geq 0$. We denote the projection $\mu(\cdot \times \{i\})$ by μ_i .

In general, an invariant measure doesn't have to be absolutely continuous with respect to the product of Lebesgue measure on \mathbb{R} and counting measure on S (in the sequel, we will simply say "with respect to Lebesgue measure"). When it is, the density of the invariant measure with respect to Lebesgue measure (which is guaranteed by the Radon–Nikodym theorem) will be called an invariant density. An invariant density ρ is defined on $\mathbb{R} \times S$, and we will usually consider the projections $(\rho_i)_{i \in S}$ that are defined on \mathbb{R} by $\rho_i(\xi) = \rho(\xi, i)$. In an abuse of terminology, we refer to these projections as invariant densities of the invariant measure. These invariant densities are, of course, elements of $L^1(\mathbb{R})$ and whenever we state a regularity property of ρ_i , we mean to say that the equivalence class ρ_i has a representative with this regularity property.

We call a point $\xi \in \mathbb{R}$ **noncritical** if $u_i(\xi) \neq 0$ for all $i \in S$. We call $\xi \in \mathbb{R}$ **uniformly critical** if $u_i(\xi) = 0$ for all $i \in S$. Throughout this paper, we assume that the set of critical points of the vector fields in D has no accumulation points. If ξ is a critical point of a vector field u_i for some $i \in S$, we write that u_i is positive to the right of ξ if there exists an open interval with left endpoint ξ on which u_i is positive. In this definition, “right” can be replaced with “left” and “positive” with “negative”.

For $i \in S$, we denote the flow function of the vector field u_i by Φ_i . Due to forward completeness of u_i , the flow function is uniquely defined for any $t \geq 0$ and for any $\eta \in \mathbb{R}$ by

$$\begin{aligned} \frac{d}{dt}\Phi_i^t(\eta) &= u_i(\Phi_i^t(\eta)), \\ \Phi_i^0(\eta) &= \eta. \end{aligned}$$

We write \mathbb{R}_+ to denote the positive real line $(0, \infty)$. For any vector $\mathbf{t} = (t_1, \dots, t_m) \in \mathbb{R}_+^m$ of times between subsequent switches, and for any corresponding vector $\mathbf{i} = (i_1, \dots, i_m) \in S^m$ of driving states, we define

$$\Phi_{\mathbf{i}}^{\mathbf{t}}(\eta) := \Phi_{i_m}^{t_m}(\Phi_{i_{m-1}}^{t_{m-1}}(\dots \Phi_{i_1}^{t_1}(\eta)) \dots)$$

as the cumulative flow along the trajectories of u_{i_1}, \dots, u_{i_m} with starting point $\eta \in \mathbb{R}$.

3. MAIN RESULTS

Let μ be an invariant measure of $(P^t)_{t \geq 0}$ that is absolutely continuous with respect to Lebesgue measure. Let $(\rho_i)_{i \in S}$ denote the invariant densities associated to μ .

3.1. Smoothness at noncritical points. If n is a positive integer, we call a function \mathcal{C}^n on a set I or at a point ξ if the function is n times continuously differentiable on I or at ξ . Being \mathcal{C}^0 means being continuous.

Theorem 1. *Let $\xi \in \mathbb{R}$ be noncritical, and assume that there exist an integer $n \geq 1$ and a closed interval I containing ξ in its interior on which all vector fields in D are \mathcal{C}^{n+1} . Then, the invariant densities $(\rho_i)_{i \in S}$ are \mathcal{C}^n at ξ .*

Remark 1. The following statement is an immediate consequence of Theorem 1: If $\xi \in \mathbb{R}$ is noncritical and if all vector fields in D are \mathcal{C}^∞ on a closed interval I containing ξ in its interior, then the invariant densities are \mathcal{C}^∞ at ξ .

Remark 2. According to [BH12, Theorem 1], the following is a sufficient condition for absolute continuity (and also for uniqueness) of the invariant measure on $\mathbb{R} \times S$, provided that an invariant measure exists:

There is a point $\xi \in \mathbb{R}$ that is not uniformly critical and that is D -accessible from any starting point $\eta \in \mathbb{R}$. We say that a point ξ is D -accessible from η if for any open interval I containing ξ , there exist a finite sequence of vector fields in D and a corresponding sequence of switching times such that some point in I can be reached from η by following the flows composed of these vector fields and switching times.

If the process X is eventually confined to a bounded subset of \mathbb{R} , existence of an invariant measure is guaranteed by the Krylov–Bogoliubov method. This is for instance the case if each vector field in D has finitely many critical points, with

the smallest critical point attracting from the left and the largest critical point attracting from the right.

A sufficient condition for existence and exponential ergodicity that can also be applied in a noncompact setting has been provided in [BLBMZ12b, Assumption 1.8]:

Let $(\Phi_i)_{i \in S}$ be the flow functions associated to the vector fields $(u_i)_{i \in S}$, and assume that Φ_i^t is globally Lipschitz continuous with Lipschitz constant L_i^t for any $i \in S$ and for any $t \geq 0$. Furthermore, suppose that

$$\alpha_i := \inf_{t > 0} \left(-\frac{\ln(L_i^t)}{t} \right)$$

is a well-defined real number for any $i \in S$. If one assumes, as we do, that the switching rates are independent of the position of X , the stochastic process A on S is Markov and has an invariant measure ν . The condition

$$(2) \quad \sum_{i \in S} \nu(\{i\}) \alpha_i > 0$$

then implies existence of an invariant measure μ for (X, A) such that the distribution of $(X, A)_t$ converges to μ in a certain Wasserstein distance as t goes to infinity.

Condition 2 can be interpreted as (X, A) contracting in mean, see [CH13, page 5]. In [CH13, Assumption 1.3], the authors state a generalization of Condition 2 for switching between members of a finite family of Markov processes. See also [YZ10, Sections 3.3 and 3.4] for criteria for recurrence and positive recurrence of the process (X, A) .

3.2. Asymptotics at critical points. Let ξ be a critical point of u_i for some $i \in S$, and assume that none of the other vector fields in D have ξ as a critical point. This assumption is made to simplify the asymptotic analysis (see [BHK⁺11, Section 2]). Without loss of generality, let $u = u_1$ and let $\xi = 0$. Recall our standing assumption that for all vector fields $u_j \in D$, the set of critical points of u_j has no accumulation point (see Section 2). Then, there is a $\delta > 0$ such that none of the vector fields in $D \setminus \{u_1\}$ have a critical point in $[0, \delta]$ and u_1 has no critical point in $(0, \delta]$. To simplify the analysis further, we assume that there is a constant $a \neq 0$ so that

$$u_1(\eta) = -a\eta + O(\eta^2)$$

as η approaches 0 from the right, i.e. u_1 behaves almost linearly near 0. The constant a can be thought of as the contraction or expansion coefficient of u_1 near 0.

If u_1 was of order $O(\eta^\alpha)$ for $\alpha < 1$, the vector field would not be Lipschitz continuous and the resulting ODE could fail to have unique solutions. If u_1 was of order $O(\eta^\alpha)$ for $\alpha > 1$, identifying the asymptotically dominant term would become more complicated.

Under these assumptions, we study the asymptotic behavior of ρ_1 as η approaches 0 from the right. Due to the symmetric nature of the problem, there is no need to investigate the case of η approaching 0 from the left separately.

In Section 6, we show that the support of the measures $(\mu_i)_{i \in S}$ can be represented as a finite union of closed intervals of positive length (see Corollary 3). Let \mathcal{I} denote the collection of these intervals. If μ is ergodic, \mathcal{I} contains only one interval.

Exactly one of the following statements holds:

- (A) 0 is the left endpoint of an open interval that does not contain any points from the support of $(\mu_i)_{i \in S}$.
- (B) 0 is contained in the interior of an interval $I \in \mathcal{I}$.
- (C) 0 is the left endpoint of an interval $I \in \mathcal{I}$.

Although these statements are not formulated in terms of the given vector fields, it is easy to see which of them holds by using the algorithm at the beginning of Section 6. In case A, ρ_1 is constantly equal to zero on an open interval with left endpoint 0. Cases B and C are more intricate and are dealt with in Theorems 2 and 3. In case C, either 0 is the right endpoint of an open interval that does not contain any points from the support of $(\mu_i)_{i \in S}$, or 0 is the right endpoint of an interval $J \in \mathcal{I}$. But if 0 is both left endpoint of an interval $I \in \mathcal{I}$ and right endpoint of an interval $J \in \mathcal{I}$, it is uniformly critical (see Section 6). Since we assume that 0 is only critical for u_1 , this second scenario cannot occur.

To illustrate cases B and C, we give two simple examples of PDMPs: one that corresponds to case B and one that corresponds to case C.

Example 1. Let D be the collection of vector fields $u_1(\eta) := -\eta$, $u_2(\eta) := 1$ and $u_3(\eta) := -1$. At any given time, the process X is either attracted to the critical point 0 or moves to the left or to the right at constant speed. The associated flow functions are $\Phi_1^t(\eta) = \eta e^{-t}$, $\Phi_2^t(\eta) = \eta + t$ and $\Phi_3^t(\eta) = \eta - t$, with global Lipschitz constants $L_1^t = e^{-t}$ and $L_2^t = L_3^t = 1$ for any $t \geq 0$. If we define $\alpha_1, \alpha_2, \alpha_3$ as in Remark 2, we have $\alpha_1 = 1$ and $\alpha_2 = \alpha_3 = 0$. Since we allow switching from any vector field to any other vector field, criterion (2) implies existence of an invariant measure. Theorem 1 in [BH12] implies that the invariant measure is unique and absolutely continuous. The projections $(\mu_i)_{i \in S}$ of this invariant measure are supported on \mathbb{R} (see Section 6). This is then an example of case B.

Example 2. Let D be the collection of vector fields $u_1(\eta) := -\eta$ and $u_2(\eta) := 1 - \eta$. The process X is alternately attracted by 0 and 1, and is eventually confined to the interval $(0, 1)$. By the Krylov–Bogoliubov theorem, $(P^t)_{t \geq 0}$ has an invariant measure. The invariant measure is unique and absolutely continuous. The support of the measures $(\mu_i)_{i \in S}$ is the closed interval $[0, 1]$, so this is an example of case C.

3.2.1. *Asymptotics for analytic vector fields.* In this subsection, we suppose in addition to the assumptions of Subsection 3.2 that all vector fields in D are analytic in an open interval around 0. To state our result on the asymptotically dominant term of ρ_1 , we introduce the function

$$\bar{\rho}(\eta) := \sum_{i>1} \lambda_{i,1} \rho_i(\eta).$$

Theorem 2. *Under the assumptions above, the following statements hold.*

- (1) Let $\lambda_1 < a$. In cases B and C, there is a constant $c > 0$ such that

$$\rho_1(\eta) = c\eta^{\frac{\lambda_1}{a}-1} + o(\eta^{\frac{\lambda_1}{a}-1})$$

as η approaches 0 from the right.

- (2) Let $\lambda_1 > a > 0$. In case B,

$$\lim_{\eta \downarrow 0} \rho_1(\eta) = \frac{\bar{\rho}(0)}{\lambda_1 - a} > 0.$$

In case C, there is a constant $c > 0$ such that

$$\rho_1(\eta) = c\eta^{\frac{\lambda_1}{a}-1} + o(\eta^{\frac{\lambda_1}{a}-1})$$

as η approaches 0 from the right.

(3) Let $\lambda_1 = a$. In case B, there is a constant $c > 0$ such that

$$\rho_1(\eta) = -c \ln(\eta) + o(\ln(\eta))$$

as η approaches 0 from the right. In case C, $\rho_1(\eta)$ converges to a positive constant as η approaches 0 from the right.

Theorem 2 is proved in Section 7. Note that in the critical case $\lambda_1 = a$, the density ρ_1 is unbounded to the right of $\eta = 0$ in case B and bounded in case C.

3.2.2. *Asymptotics in the nonanalytic case.* In the absence of analyticity, we can still recover some of the results from Theorem 2.

Theorem 3. *Under the assumptions above, without analyticity of the vector fields, the following statements hold.*

(1) Let $\lambda_1 < a$. In cases B and C, there is a constant $c > 0$ such that

$$\rho_1(\eta) = c\eta^{\frac{\lambda_1}{a}-1} + o(\eta^{\frac{\lambda_1}{a}-1})$$

as η approaches 0 from the right.

(2) Let $\lambda_1 > a > 0$. In case B,

$$\lim_{\eta \downarrow 0} \rho_1(\eta) = \frac{\bar{\rho}(0)}{\lambda_1 - a} > 0.$$

In case C,

$$\lim_{\eta \downarrow 0} \rho_1(\eta) = \frac{\bar{\rho}(0)}{\lambda_1 - a} = 0.$$

(3) Let $\lambda_1 = a$. In case B, there exist constants $c', c > 0$ such that

$$-c' \ln(\eta) \leq \rho_1(\eta) \leq -c \ln(\eta)$$

for η sufficiently small. In case C, there is a constant $c > 0$ such that

$$\rho_1(\eta) \leq -c \ln(\eta).$$

(4) Let $a < 0$, i.e. 0 is a repelling critical point of u_1 . In case B,

$$\lim_{\eta \downarrow 0} \rho_1(\eta) = \frac{\bar{\rho}(0)}{\lambda_1 - a} > 0.$$

Case C is not possible.

Theorem 3 is proved in Section 7. Theorem 3 implies the following conditions for boundedness of ρ_1 to the right of 0.

Corollary 1. (1) If $\lambda_1 < a$, ρ_1 is unbounded to the right of 0 in cases B and C.

(2) If $\lambda_1 > a > 0$, ρ_1 is bounded to the right of 0 in cases B and C.

(3) If $\lambda_1 = a$, ρ_1 is unbounded to the right of 0 in case B. In case C, our analysis is inconclusive.

Remark 3. The conditions in Corollary 1 align with intuition. If $\lambda_1 < a$, the rate of switching away from u_1 is lower than the rate at which u_1 contracts to its critical point 0. In this case, the rate at which mass accumulates in the vicinity of 0 is high, which results in a singularity of the invariant density at 0. If $\lambda_1 > a > 0$, the rate of switching away from u_1 is higher than the rate of contracting to 0. The rate at which mass accumulates at 0 is low and ρ_1 is bounded near 0 (see [BHK⁺11, Theorem 1, part c]).

4. INTEGRAL EQUATIONS FOR THE DENSITIES

In this section, we present two integral equations satisfied almost everywhere by invariant densities of $(\mathbf{P}^t)_{t \geq 0}$. Loosely stated, these equations illustrate how mass with respect to an invariant density ρ_i accumulates at a point $\eta \in \mathbb{R}$. At some point in time, there is a switch from a vector field in $D \setminus \{u_i\}$ to u_i , and the flow associated to u_i transports mass to η . In a sense, we condition on the time and nature of this last switch to u_i . The family of equations in Lemma 1 describe the mass transport for a finite history of the process. In this case, there is a positive probability of having no switch. The equation in Lemma 2 describes the transport mechanism for an infinite history. This guarantees that with probability 1, there is at least one switch.

Let μ be an absolutely continuous invariant measure of $(\mathbf{P}^t)_{t \geq 0}$, with invariant densities $(\rho_i)_{i \in S}$. Since we do not assume backward completeness of the vector fields in D , we have to be careful when studying the history of a switching trajectory. It could happen that the backward flow associated to a vector field goes off to $-\infty$ or ∞ in finite time. For any $\eta \in \mathbb{R}$ and for any $i \in S$, let $\tau_i(\eta)$ denote the supremum over the set of times $t \geq 0$ for which $t \mapsto \Phi_i^{-t}(\eta)$ is well-defined. With this definition, we introduce the shorthand

$$(3) \quad \Phi_i^t \# h(\eta) := \begin{cases} \frac{h(\Phi_i^{-t}(\eta))}{D\Phi_i^t(\Phi_i^{-t}(\eta))}, & t < \tau_i(\eta) \\ 0, & t \geq \tau_i(\eta) \end{cases}$$

for the pushforward of the function h under the flow map Φ_i^t . We think of h as a density function on the real line. Note that $D\Phi_i^t > 0$ in dimension one, so there is no need for absolute value in the denominator. Since u_i is assumed to be \mathcal{C}^1 , so is $\eta \mapsto \Phi_i^t(\eta)$, and the differential $D\Phi_i^t$ is well-defined.

Let $L_+^1(\mathbb{R})$ be the set of L^1 -functions on the real line that have a nonnegative representative. In other words, $L_+^1(\mathbb{R})$ is the space of densities for finite measures on \mathbb{R} .

For any $h \in L_+^1(\mathbb{R})$ and for any $T > 0$, define the Perron–Frobenius operators

$$(4) \quad \bar{\mathbf{P}}_i^T h(\eta) := \frac{1}{T} \int_0^T e^{-\lambda_i t} \Phi_i^t \# h(\eta) dt$$

and

$$(5) \quad \hat{\mathbf{P}}_i^T h(\eta) := \frac{1}{T} \int_0^T e^{-\lambda_i t} (T - t) \Phi_i^t \# h(\eta) dt.$$

We can now state the truncated version of the integral equation.

Lemma 1. *For any $i \in S$ and for any $T > 0$,*

$$\rho_i \equiv \bar{P}_i^T \rho_i + \sum_{j \neq i} \lambda_{j,i} \hat{P}_i^T \rho_j.$$

To state the integral equation over an infinite time horizon, we define the operators

$$\bar{P}_i h(\eta) := \int_0^\infty e^{-\lambda_i t} \Phi_i^t \# h(\eta) dt, \quad i \in S$$

for densities $h \in L_+^1(\mathbb{R})$.

Lemma 2. *For any $i \in S$,*

$$\rho_i = \sum_{j \neq i} \lambda_{j,i} \bar{P}_i \rho_j.$$

Lemmas 1 and 2 are proved in Section 8. As will become apparent from these proofs, the lemmas continue to hold if the state space \mathbb{R} of the continuous component X is replaced with a finite-dimensional smooth manifold.

5. SMOOTHNESS OF THE DENSITIES AT NONCRITICAL POINTS

In this section, we prove Theorem 1, which was stated at the beginning of Section 3. By assumption, there exist an integer $n \geq 1$ and a closed interval I , with ξ in its interior, on which all vector fields in D are \mathcal{C}^{n+1} . Since ξ is noncritical and since for each vector field in D , the set of critical points has no accumulation point, we may assume without loss of generality that I does not contain any critical points. Let $I_0 \subset I$ be another compact interval containing ξ in its interior, whose endpoints are a positive distance away from the endpoints of I . As the trajectories of the X -component of the switching process have bounded speed on compact subsets of \mathbb{R} , there is a small time $T_0 > 0$ such that $(\Phi_{\mathbf{i}}^{\mathbf{s}})^{-1}(\eta) \in I$ for any finite index sequence \mathbf{i} , any corresponding sequence of nonnegative switching times \mathbf{s} with l^1 -norm less than or equal to T_0 and for any $\eta \in I_0$.

We define the integration kernels

$$(6) \quad \mathcal{K}_i(\zeta, \eta) := \frac{\exp\left(\lambda_i \int_\eta^\zeta \frac{dx}{u_i(x)}\right)}{u_i(\eta)}$$

and

$$(7) \quad \hat{\mathcal{K}}_i^{T_0}(\zeta, \eta) := \left(T_0 + \int_\eta^\zeta \frac{dx}{u_i(x)}\right) \mathcal{K}_i(\zeta, \eta)$$

for $i \in S$ and $(\zeta, \eta) \in I \times I_0$. With these definitions, we have the following representations of $\bar{P}_i^{T_0} \rho_i$ and $\hat{P}_i^{T_0} \rho_j$. See (4) and (5) for the definitions of $\bar{P}_i^{T_0}$ and $\hat{P}_i^{T_0}$.

Lemma 3. *For any $i \in S$ and for any $\eta \in I_0$,*

$$(8) \quad \bar{P}_i^{T_0} \rho_i(\eta) = \frac{1}{T_0} \int_{\Phi_i^{-T_0}(\eta)}^\eta \rho_i(\zeta) \mathcal{K}_i(\zeta, \eta) d\zeta.$$

For any $i, j \in S$, $i \neq j$, and for any $\eta \in I_0$,

$$(9) \quad \hat{P}_i^{T_0} \rho_j(\eta) = \frac{1}{T_0} \int_{\Phi_i^{-T_0}(\eta)}^\eta \rho_j(\zeta) \hat{\mathcal{K}}_i^{T_0}(\zeta, \eta) d\zeta.$$

Our definition of T_0 implies that the interval $[\Phi_i^{-T_0}(\eta), \eta]$ (or $[\eta, \Phi_i^{-T_0}(\eta)]$ if $u_i(\xi) < 0$) is contained in I , so the integrals on the right are well-defined. Notice in particular that this reasoning still holds if u_i is not backward complete.

PROOF: Fix an $\eta \in I_0$ and recall the definition of $\Phi_i^t \# \rho_i$ in (3). Linearity of the Jacobi flow gives

$$D\Phi_i^t(\Phi_i^{-t}(\eta)) = \frac{u_i(\eta)}{u_i(\Phi_i^{-t}(\eta))},$$

hence

$$\Phi_i^t \# \rho_i(\eta) = \rho_i(\Phi_i^{-t}(\eta)) \frac{u_i(\Phi_i^{-t}(\eta))}{u_i(\eta)}$$

for any $t \in [0, T_0]$. The change of variables $\zeta = \Phi_i^{-t}(\eta)$ then yields (8). Formula (9) is proved similarly. \square

In (8) and (9), the expressions on the right still make sense if \mathcal{K}_i and $\hat{\mathcal{K}}_i^{T_0}$ are replaced with arbitrary kernels on $I \times I_0$. For any such kernel \mathcal{K} and for any $i, j \in S$, set

$$(10) \quad \mathcal{K}_i^{T_0} \rho_j(\eta) := \frac{1}{T_0} \int_{\Phi_i^{-T_0}(\eta)}^{\eta} \rho_j(\zeta) \mathcal{K}(\zeta, \eta) d\zeta.$$

The following lemma addresses regularity of the integration kernels $(\mathcal{K}_i)_{i \in S}$ and $(\hat{\mathcal{K}}_i^{T_0})_{i \in S}$.

Lemma 4. *The kernels $(\mathcal{K}_i)_{i \in S}$ and $(\hat{\mathcal{K}}_i^{T_0})_{i \in S}$ are \mathcal{C}^{n+1} on $I \times I_0$.*

PROOF: This follows from the fact that u_i is \mathcal{C}^{n+1} and nonzero on I . \square

The following lemmas illustrate the smoothening effect of the operators $(\bar{\mathcal{P}}_i^{T_0})_{i \in S}$ and $(\hat{\mathcal{P}}_i^{T_0})_{i \in S}$. We begin by showing that, away from critical points, the densities $(\rho_i)_{i \in S}$ are bounded.

Lemma 5. *The densities $(\rho_i)_{i \in S}$ are bounded on the interval I_0 .*

PROOF: Fix an $i \in S$. By Lemma 1, it is enough to show that $\bar{\mathcal{P}}_i^{T_0} \rho_i$ and $(\hat{\mathcal{P}}_i^{T_0} \rho_j)_{j \neq i}$ are bounded on I_0 . Since \mathcal{K}_i and $\hat{\mathcal{K}}_i^{T_0}$ are \mathcal{C}^1 on the compact set $I \times I_0$ (Lemma 4), they are also bounded on $I \times I_0$ by constants k_i and k_{i, T_0} . For $j \in S$, let $\|\rho_j\|_1$ denote the L^1 -norm of ρ_j on \mathbb{R} . Using (8),

$$\bar{\mathcal{P}}_i^{T_0} \rho_i(\eta) \leq \frac{k_i}{T_0} \int_{\Phi_i^{-T_0}(\eta)}^{\eta} \rho_i(\zeta) d\zeta \leq \frac{k_i \|\rho_i\|_1}{T_0}$$

for any $\eta \in I_0$.

And using (9),

$$\hat{\mathcal{P}}_i^{T_0} \rho_j(\eta) \leq \frac{k_{i, T_0}}{T_0} \int_{\Phi_i^{-T_0}(\eta)}^{\eta} \rho_j(\zeta) d\zeta \leq \frac{k_{i, T_0} \|\rho_j\|_1}{T_0}$$

for any $j \neq i, \eta \in I_0$. \square

Remark 4. In the proof of Lemma 5, we did not use any concrete information about \mathcal{K}_i or $\hat{\mathcal{K}}_i^{T_0}$ other than boundedness on $I \times I_0$. The result still holds if \mathcal{K}_i and $\hat{\mathcal{K}}_i^{T_0}$ are replaced with arbitrary kernels that are bounded on $I \times I_0$. Furthermore, the time T_0 can be replaced with any time $T \in (0, T_0)$.

The following corollary will be useful in Section 7 when we derive asymptotics for invariant densities at critical points.

Corollary 2. *Let $i \in S$ and assume that $\xi \in \mathbb{R}$ is not a critical point of u_i . Then, there is a compact interval I with ξ in its interior, such that ρ_i is bounded on I .*

In Lemma 5, we assumed that ξ is noncritical. Here, the point ξ may be critical for some of the vector fields in D , just not for the particular vector field u_i whose corresponding density function we are interested in.

PROOF: Since $u_i(\xi) \neq 0$ and since the set of critical points of u_i has no accumulation points, there is a compact interval I that has ξ in its interior, but does not contain any critical points of u_i . Let $I_0 \subset I$ be another compact interval with ξ in its interior such that the endpoints of I_0 are a positive distance away from the endpoints of I . Choose $T > 0$ so small that $\Phi_i^{-t}(\eta) \in I$ for any $\eta \in I_0$ and for any $t \in [0, T]$. Define the kernels \mathcal{K}_i and $\hat{\mathcal{K}}_i^T$ according to (6) and (7). These kernels are bounded on $I \times I_0$, and we can repeat the proof of Lemma 5 to finish the argument. \square

Let $I_1 \subset I_0$ be a compact interval that contains ξ in its interior and whose endpoints are a positive distance away from the endpoints of I_0 . Let $T_1 \in (0, T_0]$ be so small that $(\Phi_{\mathbf{i}}^s)^{-1}(\eta) \in I_0$ for any index sequence \mathbf{i} , any corresponding sequence of nonnegative switching times \mathbf{s} with l^1 -norm less than or equal to T_1 , and for any $\eta \in I_1$.

Lemma 6. *The densities $(\rho_i)_{i \in S}$ are Lipschitz continuous on I_1 .*

PROOF: Fix an $i \in S$. By Lemma 1, it is enough to show that $\bar{\mathbb{P}}_i^{T_1} \rho_i$ and $(\hat{\mathbb{P}}_i^{T_1} \rho_j)_{j \neq i}$ are Lipschitz continuous on I_1 . By Lemma 5, ρ_i is bounded on I_0 by some constant r_i . Let L be a Lipschitz constant of \mathcal{K}_i on $I \times I_0$ and let \tilde{L} be a Lipschitz constant of the flow function Φ_i on $[-T_1, 0] \times I_1$. The constant k_i is defined as in the proof of Lemma 5 and k_{i, T_1} is defined in analogy to k_{i, T_0} . Fix two points $\eta, \vartheta \in I_1$. As $\Phi_i^{-T_1}(\eta)$ and $\Phi_i^{-T_1}(\vartheta)$ are both contained in I_0 , we obtain the estimate

$$\begin{aligned} & |\bar{\mathbb{P}}_i^{T_1} \rho_i(\eta) - \bar{\mathbb{P}}_i^{T_1} \rho_i(\vartheta)| \\ &= \frac{1}{T_1} \left| \int_{\Phi_i^{-T_1}(\eta)}^{\eta} \rho_i(\zeta) \mathcal{K}_i(\zeta, \eta) d\zeta - \int_{\Phi_i^{-T_1}(\vartheta)}^{\vartheta} \rho_i(\zeta) \mathcal{K}_i(\zeta, \vartheta) d\zeta \right| \\ &\leq \frac{1}{T_1} \left(\left| \int_{\Phi_i^{-T_1}(\eta)}^{\Phi_i^{-T_1}(\vartheta)} \rho_i(\zeta) \mathcal{K}_i(\zeta, \eta) d\zeta \right| + \left| \int_{\eta}^{\vartheta} \rho_i(\zeta) \mathcal{K}_i(\zeta, \vartheta) d\zeta \right| \right. \\ &\quad \left. + \left| \int_{\Phi_i^{-T_1}(\vartheta)}^{\eta} \rho_i(\zeta) (\mathcal{K}_i(\zeta, \eta) - \mathcal{K}_i(\zeta, \vartheta)) d\zeta \right| \right) \\ &\leq \|\vartheta - \eta\| \frac{1}{T_1} (r_i k_i (1 + \tilde{L}) + L \|\rho_i\|_1). \end{aligned}$$

Let \hat{L} be a Lipschitz constant of $\hat{\mathcal{K}}_i^{T_1}$ on $I \times I_0$. For a fixed $j \neq i$, the density ρ_j is bounded on I_0 by a constant r_j , and

$$(11) \quad |\hat{\mathbb{P}}_i^{T_1} \rho_j(\eta) - \hat{\mathbb{P}}_i^{T_1} \rho_j(\vartheta)| \leq \|\vartheta - \eta\| \frac{1}{T_1} (r_j k_{i, T_1} (1 + \tilde{L}) + \hat{L} \|\rho_j\|_1).$$

\square

Remark 5. Lemma 6 continues to hold if \mathcal{K}_i and $\hat{\mathcal{K}}_i^{T_1}$ are replaced with arbitrary kernels that are Lipschitz continuous on $I \times I_0$ and if T_1 is replaced with an arbitrary time $T \in (0, T_1)$.

Remark 6. Lemma 6 implies the following: If an open interval I does not contain any critical points, then all invariant densities ρ_i are Lipschitz continuous on I . Slightly modifying the proof of Lemma 6, one can show a related statement:

If an open interval I does not contain any critical points of a particular vector field u_i (but possibly critical points of other vector fields), the invariant density ρ_i is continuous on I .

Notice that we can only guarantee the continuity of ρ_i , not its Lipschitz continuity. Since we allow for critical points of the other vector fields $(u_j)_{j \neq i}$ on I , we also can no longer ascertain boundedness of the corresponding densities $(\rho_j)_{j \neq i}$. Instead of (11), we obtain the weaker estimate

$$\begin{aligned} |\hat{\rho}_i^{T_1} \rho_j(\eta) - \hat{\rho}_i^{T_1} \rho_j(\vartheta)| &\leq \frac{k_{i,T_1}}{T_1} \left(\left| \int_{\Phi_i^{-T_1}(\vartheta)}^{\Phi_i^{-T_1}(\eta)} \rho_j(\zeta) d\zeta \right| + \left| \int_{\vartheta}^{\eta} \rho_j(\zeta) d\zeta \right| \right) \\ &\quad + \frac{\|\rho_j\|_1}{T_1} \hat{L} |\vartheta - \eta|. \end{aligned}$$

Lemma 7 illustrates the actual smoothening mechanism.

Lemma 7. *For any integer $k \in \{0, \dots, n-1\}$, the following statement holds:*

If the densities $(\rho_i)_{i \in S}$ are \mathcal{C}^k on a compact interval $\tilde{I} \subset I_1$ that contains ξ in its interior, there exist a compact interval $\tilde{I}' \subset \tilde{I}$ with ξ in its interior and a time $T \in (0, T_1]$ such that for any \mathcal{C}^{k+2} -kernel \mathcal{K} on $I \times I_0$, the functions $(\mathcal{K}_i^T \rho_j)_{i,j \in S}$ are \mathcal{C}^{k+1} on \tilde{I}' .

Recall that we defined $\mathcal{K}_i^T \rho_j$ in (10).

PROOF: We prove Lemma 7 by induction on k . In the base case, assume that the densities $(\rho_i)_{i \in S}$ are continuous on $\tilde{I} \subset I_1$. Let $\tilde{I}' \subset \tilde{I}$ be a compact interval that contains ξ in its interior and whose endpoints are a positive distance away from the endpoints of \tilde{I} . Let $T \in (0, T_1]$ be so small that $(\Phi_i^{\mathbf{s}})^{-1}(\eta) \in \tilde{I}$ for any index sequence \mathbf{i} , any corresponding sequences of nonnegative switching times \mathbf{s} with l^1 -norm less than or equal to T , and for any $\eta \in \tilde{I}'$. For any \mathcal{C}^2 -kernel \mathcal{K} on $I \times I_0$, for any $\eta \in \tilde{I}'$ and for any $i, j \in S$,

$$\begin{aligned} \frac{d}{d\eta} \mathcal{K}_i^T \rho_j(\eta) &= \frac{1}{T} \left(\rho_j(\eta) \mathcal{K}(\eta, \eta) - \rho_j(\Phi_i^{-T}(\eta)) \mathcal{K}(\Phi_i^{-T}(\eta), \eta) \frac{d}{d\eta} \Phi_i^{-T}(\eta) \right) \\ &\quad + \frac{1}{T} \int_{\Phi_i^{-T}(\eta)}^{\eta} \rho_j(\zeta) \partial_2 \mathcal{K}(\zeta, \eta) d\zeta \\ (12) \quad &= \frac{1}{T} \left(\rho_j(\eta) \mathcal{K}(\eta, \eta) - \rho_j(\Phi_i^{-T}(\eta)) \mathcal{K}(\Phi_i^{-T}(\eta), \eta) \frac{d}{d\eta} \Phi_i^{-T}(\eta) \right) \\ &\quad + (\partial_2 \mathcal{K})_i^T \rho_j(\eta). \end{aligned}$$

Here, $\partial_2 \mathcal{K}$ denotes the partial derivative of \mathcal{K} with respect to its second component. Since ρ_j is assumed to be \mathcal{C}^0 on \tilde{I} , since \mathcal{K} is \mathcal{C}^2 on $I \times I_0$ and since u_i is \mathcal{C}^{n+1} on I , the first term in (12) is \mathcal{C}^0 on \tilde{I}' .

It remains to show that $(\partial_2 \mathcal{K})_i^T \rho_j$ is \mathcal{C}^0 , but this follows along the lines of Lemma 6, keeping in mind that $\partial_2 \mathcal{K}$ is \mathcal{C}^1 and hence Lipschitz continuous on $I \times I_0$

and that $T \leq T_1$ (see Remark 5). Since $\frac{d}{d\eta} \mathcal{K}_i^T \rho_j(\eta)$ is \mathcal{C}^0 on \tilde{I}' , it follows that $\mathcal{K}_i^T \rho_j$ is \mathcal{C}^1 on \tilde{I}' . This completes the base case.

In the induction step, let k be a fixed integer in $\{1, \dots, n-1\}$ and assume that the statement holds for $k-1$. Assume that the densities $(\rho_i)_{i \in S}$ are \mathcal{C}^k on $\tilde{I} \subset I_1$. The densities $(\rho_i)_{i \in S}$ are then also \mathcal{C}^{k-1} on \tilde{I} . By induction hypothesis, there exist a compact interval $\tilde{I}' \subset \tilde{I}$ with ξ in its interior and a time $T \in (0, T_1]$ such that for any \mathcal{C}^{k+1} -kernel \mathcal{K} on $I \times I_0$, the functions $(\mathcal{K}_i^T \rho_j)_{i,j \in S}$ are \mathcal{C}^k on \tilde{I}' . Without loss of generality, we can assume that the endpoints of \tilde{I}' are a positive distance away from the endpoints of \tilde{I} and that T is so small that $(\Phi_{\mathbf{i}}^s)^{-1}(\eta) \in \tilde{I}$ for any index sequence \mathbf{i} , any corresponding sequence of nonnegative switching times \mathbf{s} with l^1 -norm less than or equal to T , and for any $\eta \in \tilde{I}'$. Let \mathcal{K} be a \mathcal{C}^{k+2} -kernel on $I \times I_0$. Then, (12) holds for any $\eta \in \tilde{I}'$ and for any $i, j \in S$.

Since ρ_j is by assumption \mathcal{C}^k on \tilde{I} , since \mathcal{K} is \mathcal{C}^{k+2} on $I \times I_0$ and since u_i is \mathcal{C}^{n+1} on I , the first term in (12) is \mathcal{C}^k on \tilde{I}' . In addition, $\partial_2 \mathcal{K}$ is a \mathcal{C}^{k+1} -kernel on $I \times I_0$. By induction hypothesis, $(\partial_2 \mathcal{K})_i^T \rho_j$ is \mathcal{C}^k on \tilde{I}' , so $\frac{d}{d\eta} \mathcal{K}_i^T \rho_j(\eta)$ is \mathcal{C}^k on \tilde{I}' . From this, it follows that $\mathcal{K}_i^T \rho_j$ is \mathcal{C}^{k+1} on \tilde{I}' . \square

PROOF OF THEOREM 1: In order to prove Theorem 1, it suffices to show the following statement: For any integer $k \in \{0, \dots, n\}$, there is a compact interval I_{k+1} with ξ in its interior such that the densities $(\rho_i)_{i \in S}$ are \mathcal{C}^k on I_{k+1} .

We prove this statement by induction on k . By Lemma 6, the densities $(\rho_i)_{i \in S}$ are Lipschitz continuous on I_1 . This takes care of the base case. In the induction step, let k be an integer in $\{1, \dots, n\}$ and assume that the densities $(\rho_i)_{i \in S}$ are \mathcal{C}^{k-1} on a compact interval $I_k \subset I_1$ with ξ in its interior. By Lemma 7, there exist a compact interval $I_{k+1} \subset I_k$ with ξ in its interior and a time $T \in (0, T_1]$ such that for any \mathcal{C}^{k+1} -kernel \mathcal{K} on $I \times I_0$, the functions $(\mathcal{K}_i^T \rho_j)_{i,j \in S}$ are \mathcal{C}^k on I_{k+1} .

Fixing an $i \in S$, Lemma 7 applied to the integration kernel \mathcal{K}_i yields that $\bar{\mathcal{P}}_i^T \rho_i$ is \mathcal{C}^k on I_{k+1} . Applying Lemma 7 to $\hat{\mathcal{K}}_i^T$ yields that for any $j \neq i$, $\hat{\mathcal{P}}_i^T \rho_j$ is \mathcal{C}^k on I_{k+1} . By Lemma 1, ρ_i is \mathcal{C}^k on I_{k+1} . \square

6. THE SUPPORT OF INVARIANT MEASURES ON THE REAL LINE

Let μ be an invariant measure of the Markov semigroup $(\mathbf{P}^t)_{t \geq 0}$. In this section, we describe the support of the measures $(\mu_i)_{i \in S}$, which are measures on the real line. Recall that a point $\xi \in \mathbb{R}$ lies in the support of μ_i if and only if $\mu_i(U) > 0$ for any open neighborhood U of ξ .

We say that a point $\xi \in \mathbb{R}$ is D -reachable from a point $\eta \in \mathbb{R}$ if there exist a finite index sequence \mathbf{i} and a corresponding sequence of nonnegative switching times \mathbf{t} such that

$$\Phi_{\mathbf{i}}^{\mathbf{t}}(\eta) = \xi.$$

For any $\xi \in \mathbb{R}$, we define $L(\xi)$ as the set of points that are D -reachable from ξ .

We call a point $\xi \in \mathbb{R}$ D -accessible from $\eta \in \mathbb{R}$ if for any open neighborhood U of ξ there exist a finite index sequence \mathbf{i} and a corresponding sequence of nonnegative switching times \mathbf{t} such that

$$\Phi_{\mathbf{i}}^{\mathbf{t}}(\eta) \in U.$$

Let L denote the set of points on the real line that are D -accessible from any point in \mathbb{R} .

6.1. **Minimal invariant sets.** A nonempty set $I \subset \mathbb{R}$ is called invariant if

$$\Phi_1^{\mathbf{t}}(\xi) \in I$$

for any $\xi \in I$, any finite index sequence \mathbf{i} and any corresponding sequence of nonnegative switching times \mathbf{t} . A minimal invariant set is an invariant set for which any nonempty strict subset is not invariant. Alternatively, a minimal invariant set is a nonempty set I with the property that

$$(13) \quad L(\xi) = I$$

for any $\xi \in I$.

The following algorithm yields exactly the minimal invariant sets of our switching system.

- (1) Mark $-\infty$ with the label “l” and mark $+\infty$ with the label “r”.
- (2) Mark those critical points where all vector fields in D are nonnegative with an “l” and mark those critical points where all vector fields in D are nonpositive with an “r”. If a critical point has both labels “l” and “r”, it is uniformly critical. All uniformly critical points form minimal invariant sets.
- (3) Consider all points, including $-\infty$, with the label “l”. This includes those points that carry both labels. As $+\infty$ doesn’t have label “l”, each of these points has a closest labeled point to its right. If this point has label “r”, the open, possibly infinite, interval with the “l”-labeled and the “r”-labeled points as its endpoints is a candidate for a minimal invariant set. It is indeed a minimal invariant set if and only if it contains two not necessarily distinct points ξ and η for which there are vector fields $u, v \in D$ with $u(\xi) > 0$ and $v(\eta) < 0$.

Proposition 1. *The algorithm above characterizes the minimal invariant sets of the switching system completely. Minimal invariant sets are thus either open intervals or point sets with exactly one element.*

PROOF: We first show that any set identified by the algorithm as a minimal invariant set is indeed a minimal invariant set. Let \mathcal{S} be a set identified by the algorithm as a minimal invariant set. Then, either $\mathcal{S} = \{\xi\}$, where ξ is a uniformly critical point, or \mathcal{S} is an open interval (l, r) , where $l < r$ are elements of the extended real line such that

- (a) $l = -\infty$ or $u_i(l) \geq 0$ for any $i \in S$
- (b) $r = \infty$ or $u_i(r) \leq 0$ for any $i \in S$
- (c) for any critical point ξ in (l, r) there exist indices $i, j \in S$ with $u_i(\xi) < 0 < u_j(\xi)$
- (d) if there are no critical points in (l, r) , there are at least points $\xi, \eta \in (l, r)$ and indices $i, j \in S$ with $u_i(\xi) < 0 < u_j(\eta)$.

If $\mathcal{S} = \{\xi\}$, it is clear that \mathcal{S} is a minimal invariant set: The only strict subset of \mathcal{S} is the empty set, and \mathcal{S} is invariant because ξ is uniformly critical.

If $\mathcal{S} = (l, r)$, no switching trajectory starting in \mathcal{S} can get to the left of l or to the right of r . This is obvious if $l = -\infty$ or $r = \infty$. If l or r are finite, it is guaranteed by Conditions a and b, respectively. Hence, \mathcal{S} is invariant. Next, we show that \mathcal{S} is also minimal. Assume that \mathcal{S} is not minimal. Then, there is a nonempty strict subset \mathcal{R} of \mathcal{S} that is invariant. In addition, there is a point $\zeta \in \mathcal{S}$ with $u_i(\zeta) \leq 0$ for any $i \in S$. To see this, fix a point $\eta \in \mathcal{S} \setminus \mathcal{R}$ and a point $\xi \in \mathcal{R}$. We can assume

without loss of generality that $\eta > \xi$. Since \mathcal{R} is invariant, η is not D -reachable from ξ . Thus, there is a point $\zeta \in [\xi, \eta]$ with $u_i(\zeta) \leq 0$ for any $i \in S$.

In light of Condition c, ζ is not critical. On the other hand, Condition d implies that there exist a $\tilde{\zeta} \in \mathcal{S}$ and a $j \in S$ with $u_j(\tilde{\zeta}) > 0$. Assume without loss of generality that $\tilde{\zeta} > \zeta$, and let

$$\hat{\zeta} := \sup\{\theta \in [\zeta, \tilde{\zeta}] : u_i(\theta) < 0 \forall i \in S\}.$$

The point $\hat{\zeta}$ is a critical point in \mathcal{S} with $u_i(\hat{\zeta}) \leq 0$ for any $i \in S$. This violates Condition c.

Conversely, let I be a minimal invariant set. We need to show that the algorithm correctly identifies I as a minimal invariant set. Due to the minimality assumption, I is an interval. If I contains exactly one point, this point is uniformly critical, for otherwise, I would not be invariant.

If I has at least two elements, it is an interval with distinct endpoints. We now show that if an endpoint of I is finite, it must be a critical point. Let ξ be a finite endpoint of I , say its left endpoint, and assume that ξ is noncritical. Since I is invariant, $u_i(\xi) > 0$ for any $i \in S$. By continuity of the vector fields, there is an $\epsilon > 0$ such that $u_i(\eta) > 0$ for any $i \in S$ and for any $\eta \in [\xi, \xi + \epsilon]$. By choosing ϵ sufficiently small, we can then ensure that $I \setminus [\xi, \xi + \epsilon]$ is a nonempty strict subset of I that is invariant. This contradicts the minimality assumption on I .

Invariance of I also implies that the endpoints of I are not D -reachable from a starting point in the interior of I . Hence, I is an open interval (l, r) , where l and r may be finite or infinite.

It remains to show that Conditions c and d are satisfied. Let $\xi \in I$ be a critical point. If $u_i(\xi) \geq 0$ for any $i \in S$, the interval $(\xi, r) \subset (l, r)$ is invariant as well which is a contradiction. Similarly, $u_i(\xi)$ cannot be nonpositive for all $i \in S$, so we can find $i, j \in S$ with $u_i(\xi) < 0 < u_j(\xi)$. To show that Condition d holds, assume that $u_i(\eta) \geq 0$ for all $\eta \in I$ and for all $i \in S$. For $\xi \in I$, the interval (ξ, r) is invariant, which contradicts the minimality assumption. \square

Proposition 2. *Minimal invariant sets are pairwise disjoint.*

PROOF: Let I and J be minimal invariant sets with $I \cap J \neq \emptyset$. As the intersection of invariant sets, $I \cap J$ is invariant. Since I and J are minimal, it follows that $I = I \cap J = J$. \square

6.2. How minimal invariant sets relate to the support of invariant measures. We begin by relating invariant measures of the global dynamics on $\mathbb{R} \times S$ to invariant measures of the switching dynamics confined to a minimal invariant set.

Let $I \subset \mathbb{R}$ be a minimal invariant set. On $I \times S$, we define the semigroup $(\mathbf{p}^t)_{t \geq 0}$ by

$$\mathbf{p}_{\xi, i}^t(E \times \{j\}) := \mathbf{P}_{\xi, i}^t(E \times \{j\})$$

for any $(\xi, i) \in I \times S$ and for any set E in the Borel σ -algebra on I . Hence, $(\mathbf{p}^t)_{t \geq 0}$ can be thought of as the restriction of $(\mathbf{P}^t)_{t \geq 0}$ to $I \times S$. It is well-defined because I is invariant.

Proposition 3. *There is a one-to-one correspondence between invariant measures of $(\mathbf{p}^t)_{t \geq 0}$ and those invariant measures of $(\mathbf{P}^t)_{t \geq 0}$ that assign mass 1 to $I \times S$.*

This is easy to see. We omit the proof.

Next, we show that the support of the measure $\mu(\cdot \times S)$ does not contain points outside of the closure of minimal invariant sets.

Proposition 4. *Let $\xi \in \mathbb{R}$ be a point that is not contained in the closure of a minimal invariant set. Then, ξ is not contained in the support of $\mu(\cdot \times S)$.*

SKETCH OF PROOF: We record two statements without proofs.

First, there exist an open neighborhood U of ξ , an open set $V \subset \mathbb{R}$, a time $T > 0$, a positive integer m , an index sequence $\mathbf{i} = (i_1, \dots, i_m)$ of length m and an open subset Δ of the simplex

$$\left\{ \mathbf{s} \in (0, \infty)^{m-1} : \sum_{l=1}^{m-1} s_l < T \right\}$$

such that $\Phi_{\mathbf{i}}^{(\mathbf{s}, T - \sum_{l=1}^{m-1} s_l)}(\eta) \in V$ for any $\eta \in U, \mathbf{s} \in \Delta$, and such that $U \cap L(\vartheta) = \emptyset$ for any $\vartheta \in V$.

Second, there are constants $\epsilon', c' > 0$ and an open set $U' \subset U$ such that

$$\inf_{\eta \in U'; i, j \in S} \mathbf{P}_{\eta, i}^{\epsilon'}(U \times \{j\}) \geq c'.$$

The second statement is shown in [BLBMZ12a]. It is an immediate consequence of the fact that the speed of the process X is bounded on bounded sets.

To derive a contradiction, we assume that ξ belongs to the support of $\mu(\cdot \times S)$. As U' is an open neighborhood of ξ ,

$$c := \mu(U' \times S) > 0.$$

Therefore,

$$\begin{aligned} \sum_{i \in S} \int_{U'} \mathbf{P}_{\theta, i}^{T+\epsilon'}(V \times S) \mu_i(d\theta) &= \sum_{i \in S} \int_{U'} \sum_{l \in S} \int_{\mathbb{R}} \mathbf{P}_{\eta, i}^T(V \times S) \mathbf{P}_{\theta, i}^{\epsilon'}(d\eta \times \{l\}) \mu_i(d\theta) \\ &\geq \sum_{i \in S} \int_{U'} \int_U \mathbf{P}_{\eta, i_1}^T(V \times S) \mathbf{P}_{\theta, i}^{\epsilon'}(d\eta \times \{i_1\}) \mu_i(d\theta) \\ &\geq cc' \inf_{\eta \in U} \mathbf{P}_{\eta, i_1}^T(V \times S). \end{aligned}$$

Next, we show that $\inf_{\eta \in U} \mathbf{P}_{\eta, i_1}^T(V \times S) > 0$. Fix a point $\eta \in U$ and let $C_{\mathbf{i}}$ denote the event that the driving vector fields up to time T appear in the order given by \mathbf{i} . Let $\mathbf{P}_{\mathbf{i}}$ be the probability that the first m driving vector fields appear in the order given by \mathbf{i} , conditioned on u_{i_1} being the first driving vector field. Similarly to Lemma 17 in Section 8, we have

$$\begin{aligned} \mathbf{P}_{\eta, i_1}^T(V \times S) &\geq \mathbf{P}_{\eta, i_1}(X_T \in V, C_{\mathbf{i}}) \\ (14) \quad &\geq \mathbf{P}_{\mathbf{i}} \int_{\Delta} \prod_{l=1}^{m-1} \lambda_{i_l} e^{-\lambda_{i_l} s_l} e^{-\lambda_{i_m}(T - (s_1 + \dots + s_{m-1}))} ds. \end{aligned}$$

The term in (14) is strictly positive and does not depend on η . We conclude that

$$a := \sum_{i \in S} \int_{U'} \mathbf{P}_{\theta, i}^{T+\epsilon'}(V \times S) \mu_i(d\theta) > 0.$$

Hence, there is a positive integer N with $Na > 1$. Let $\mu\mathbf{P}$ denote the distribution of the Markov process (X, A) with initial distribution μ . For $0 \leq k \leq N-1$, define the event

$$E_k := \{X_{k(T+\epsilon')} \in U', X_{(k+1)(T+\epsilon')} \in V, X_{j(T+\epsilon')} \in U'^c \text{ for } k+2 \leq j \leq N\}.$$

Since the sets $(E_k)_{0 \leq k \leq N-1}$ are pairwise disjoint,

$$\mu\mathbf{P}(X_{N(T+\epsilon')} \in U'^c) \geq \sum_{k=0}^{N-1} \mu\mathbf{P}(E_k).$$

Since U' cannot be reached from any point in V , we have

$$E_k = \{X_{k(T+\epsilon')} \in U', X_{(k+1)(T+\epsilon')} \in V\}.$$

Then, for $0 \leq k \leq N-1$,

$$\mu\mathbf{P}(E_k) = \sum_{i \in S} \int_{\mathbb{R}} \sum_{j \in S} \int_{U'} \mathbf{P}_{\eta, j}^{T+\epsilon'}(V \times S) \mathbf{P}_{\theta, i}^{k(T+\epsilon')}(d\eta \times \{j\}) \mu_i(d\theta) = a$$

because μ is an invariant measure. We infer that

$$\mu\mathbf{P}(X_{N(T+\epsilon')} \in U'^c) \geq Na > 1,$$

which is impossible. Hence, ξ is not contained in the support of $\mu(\cdot \times S)$. \square

In Proposition 3, we saw that invariant measures on minimal invariant sets correspond to invariant measures on \mathbb{R} that are supported on a minimal invariant set. In the following proposition, we show uniqueness of the invariant measure on a given minimal invariant set.

Proposition 5. *Any minimal invariant set admits at most one invariant measure.*

PROOF: Let I be a minimal invariant set. If $I = \{\xi\}$ for some uniformly critical point ξ , uniqueness of the invariant measure is clear.

If I is an open interval, it does not contain any uniformly critical points by Proposition 1. By the alternative characterization of minimal invariant sets in (13), $I = L(\eta)$ for any $\eta \in I$. Thus, any point in I is D -reachable from all starting points in I . By [BH12, Theorem 1], this implies uniqueness of the invariant measure of the restricted semigroup $(\mathbf{p}^t)_{t \geq 0}$. \square

Now, assume that the invariant measure μ is ergodic. If I is a minimal invariant set, ergodicity of μ implies that $\mu(I \times S)$ is either 0 or 1. It is then natural to ask whether we can assign a unique minimal invariant set I to μ for which $\mu(I \times S) = 1$. The following proposition shows that this can be done.

Proposition 6. *If μ is ergodic, there is a unique minimal invariant set I with $\mu(I \times S) = 1$.*

PROOF: Let us begin by showing that such a minimal invariant set exists. Since μ is ergodic, it is enough to show that $\mu(I \times S) > 0$ for some minimal invariant set I .

We denote the set of points not contained in the closure of a minimal invariant set by \mathcal{T} . According to Proposition 4, the intersection of \mathcal{T} and of the support of $\mu(\cdot \times S)$ is empty, so there is a point $\xi \in \mathcal{T}^c$ that also lies in the support of $\mu(\cdot \times S)$. As $\xi \in \mathcal{T}^c$, there is a minimal invariant set I whose closure contains ξ . We distinguish between several cases.

First, assume that $I = \{\xi\}$. Then, ξ is uniformly critical and may or may not be an endpoint of one or two additional minimal invariant sets. If there are no minimal invariant sets adjacent to $\{\xi\}$, we can find an open neighborhood U of ξ such that $U \setminus \{\xi\} \subset \mathcal{T}$. Since the complement of the support of $\mu(\cdot \times S)$ has measure 0, it follows that $\mu(U \setminus \{\xi\} \times S) = 0$. Therefore, $\mu(\{\xi\} \times S) > 0$.

If there is at least one open minimal invariant set adjacent to $\{\xi\}$, we have $\mu(\{\xi\} \times S) > 0$, or at least one of the adjacent minimal invariant sets has strictly positive $\mu(\cdot \times S)$ -measure.

Now, assume that $I = (l, r)$. If $\xi \in I$, it is immediate from the definition of the support that $\mu(I \times S) > 0$. If ξ is an endpoint of I , assume without loss of generality that $\xi = l$. We have already dealt with the case where ξ is uniformly critical. If ξ is critical but not uniformly critical, we still have $\mu(\{\xi\} \times S) > 0$ or $\mu(I \times S) > 0$ or $\mu(J \times S) > 0$, provided that J is an open minimal invariant set with ξ as its right endpoint. We only need to exclude the case that $\mu(\{\xi\} \times S) > 0$. This can be done similarly to the proof of Proposition 4.

Uniqueness of the minimal invariant set follows from Proposition 2. \square

Proposition 7. *If μ is ergodic, there is a unique minimal invariant set I such that the support of the measures $(\mu_i)_{i \in S}$ equals the closure of I .*

PROOF: Let I be the unique minimal invariant set for which $\mu(I \times S) = 1$ and whose existence is postulated in Proposition 6. By characterization (13) of minimal invariant sets, every point in I is D -reachable from any other point in I . By [BH12, Lemma 6], I is then contained in the support of $(\mu_i)_{i \in S}$. With $\mu(I \times S) = 1$, the statement follows. \square

Corollary 3. *Let μ be an invariant measure, not necessarily ergodic. Then, there exist minimal invariant sets I_1, \dots, I_N such that the support of μ_j equals the closure of $\bigcup_{i=1}^N I_i$ for any $j \in S$. If μ is absolutely continuous, each of the minimal invariant sets I_i is an open interval.*

PROOF: This follows from Proposition 7, the Birkhoff Ergodic Theorem and the Ergodic Decomposition Theorem. See [Hai06] for discussion adapted precisely to this setting and [CFS82] for more general considerations. \square

7. ASYMPTOTICS FOR THE DENSITIES AT CRITICAL POINTS

7.1. Asymptotic analysis for nonanalytic vector fields. In this subsection, we prove Theorem 3. In both cases B and C, there is an open interval I with left endpoint 0 such that

$$\rho_i(\eta) > 0$$

for any $\eta \in I$ and for any $i \in S$. This is a consequence of the following lemma.

Lemma 8. *Let I be an open interval that is contained in the support of $(\mu_i)_{i \in S}$. If the vector field u_i does not have any critical points in I , then $\rho_i(\eta) > 0$ for any $\eta \in I$.*

PROOF: Fix a point $\eta \in I$. Let \tilde{I} be a closed subinterval of I , with η contained in the interior of \tilde{I} . Let $T > 0$ be so small that $\Phi_{\mathbf{i}}^{\mathbf{s}}(\eta) \in \tilde{I}$ for any finite index sequence \mathbf{i} and any corresponding sequence of switching times \mathbf{s} with $\|\mathbf{s}\|_1 \leq T$.

Since u_i does not have any critical points in I , the function

$$\zeta \mapsto \exp\left(-\lambda_i \int_{\zeta}^{\eta} \frac{dx}{u_i(x)}\right)$$

is bounded below on $[\Phi_i^{-T}(\eta), \eta]$ by a constant $c > 0$. Using Lemma 1 and (8), we obtain the estimate

$$\begin{aligned} \rho_i(\eta) &\geq \frac{1}{u_i(\eta)} \frac{1}{T} \int_{\Phi_i^{-T}(\eta)}^{\eta} \rho_i(\zeta) \exp\left(-\lambda_i \int_{\zeta}^{\eta} \frac{dx}{u_i(x)}\right) d\zeta \\ &\geq \frac{c}{T} \frac{1}{|u_i(\eta)|} \left| \int_{\Phi_i^{-T}(\eta)}^{\eta} \rho_i(\zeta) d\zeta \right| \\ (15) \quad &= \frac{c}{T} \frac{1}{|u_i(\eta)|} \mu_i((\Phi_i^{-T}(\eta), \eta)) > 0. \end{aligned}$$

For (15), we used that $(\Phi_i^{-T}(\eta), \eta)$ is contained in the support of μ_i . \square

Let $\delta > 0$ be so small that none of the vector fields $(u_i)_{i>1}$ have a critical point in $[0, \delta]$, and that u_1 has no critical point in $(0, \delta]$. Let $a > 0$. The vector field u_1 is then strictly negative on $(0, \delta]$.

For $\eta \in (0, \delta)$, define $\vartheta := \lim_{t \rightarrow \tau_1(\eta)} \Phi_1^{-t}(\eta)$, where $\tau_1(\eta)$ was introduced in Section 4. This limit is independent of the concrete choice of η . By Lemma 8, there is a constant $c > 0$ such that $\bar{\rho}(\eta) \geq c$ for any $\eta \in [\frac{\delta}{2}, \delta]$. In case B, we can even assume that $\bar{\rho}(\eta) \geq c$ for any $\eta \in [0, \delta]$. And by Remark 6, $\bar{\rho}$ is continuous on $[0, \delta]$, which implies that $\bar{\rho}$ is bounded from above on $[0, \delta]$ by some constant $\bar{\rho}_{\infty}$.

Set

$$r(\eta) := -\frac{1}{u_1(\eta)} - \frac{1}{a\eta}, \quad \eta \in (0, \vartheta).$$

It is not hard to see that $r(\eta)$ is bounded on $(0, \delta]$ by a constant $r_{\infty} > 0$. Furthermore, as $u_1 < 0$ on $(0, \vartheta)$, we have $r(\eta) \geq -\frac{1}{a\eta}$ for any $\eta \in (0, \vartheta)$.

For $\eta, \zeta \in [0, \vartheta]$, define

$$E(\eta, \zeta) := \exp\left(-\lambda_1 \int_{\eta}^{\zeta} r(x) dx\right).$$

Lemma 9. *The function $\zeta \mapsto \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta)$ is integrable on (δ, ϑ) for any $\eta \in [0, \delta]$.*

PROOF: For $\eta \in [0, \delta]$ and $\zeta \in (\delta, \vartheta)$,

$$\begin{aligned} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta) &= \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \delta) E(\delta, \zeta) \\ &\leq \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) e^{\lambda_1 \delta r_{\infty}} \exp\left(\lambda_1 \int_{\delta}^{\zeta} \frac{dx}{ax}\right) \\ &= \bar{\rho}(\zeta) e^{\lambda_1 \delta r_{\infty}} \delta^{-\frac{\lambda_1}{a}}. \end{aligned}$$

The fact that $\bar{\rho}$ is integrable implies the statement. \square

In analogy to Lemma 3, we have the following representation for ρ_1 .

Lemma 10. *For any $\eta \in (0, \delta)$,*

$$\rho_1(\eta) = \left(\frac{\eta^{\frac{\lambda_1}{a}-1}}{a} + r(\eta)\eta^{\frac{\lambda_1}{a}} \right) \int_{\eta}^{\vartheta} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta.$$

PROOF: Fix an $\eta \in (0, \delta)$. Using Lemma 2 and the change of variables $\zeta = \Phi_1^{-t}(\eta)$, we obtain

$$(16) \quad \rho_1(\eta) = -\frac{1}{u_1(\eta)} \int_{\eta}^{\vartheta} \bar{\rho}(\zeta) \exp\left(\lambda_1 \int_{\eta}^{\zeta} \frac{dx}{u_1(x)}\right) d\zeta.$$

Since

$$\exp\left(\lambda_1 \int_{\eta}^{\zeta} \frac{dx}{u_1(x)}\right) = \exp\left(-\lambda_1 \int_{\eta}^{\zeta} \frac{dx}{ax}\right) E(\eta, \zeta) = \eta^{\frac{\lambda_1}{a}} \zeta^{-\frac{\lambda_1}{a}} E(\eta, \zeta)$$

for any $\zeta \in (\eta, \vartheta)$, and since $\zeta \mapsto \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta)$ is integrable by Lemma 9, the statement follows. \square

PROOF OF THEOREM 3: Fix an $\eta \in (0, \delta)$. Throughout the proof, we work with the formula for ρ_1 provided in Lemma 10.

First, let $\lambda_1 < a$. Observe that $\zeta \mapsto \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(0, \zeta)$ is integrable on $(0, \delta)$ because

$$\zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(0, \zeta) \leq \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}_{\infty} e^{\lambda_1 \delta r_{\infty}}.$$

Together with Lemma 9, we see that this function is integrable on $(0, \vartheta)$, which implies that

$$\lim_{\eta \downarrow 0} \left(\int_{\eta}^{\vartheta} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta \right) = \int_0^{\vartheta} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(0, \zeta) d\zeta < \infty$$

by dominated convergence. In addition,

$$\begin{aligned} \int_0^{\vartheta} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(0, \zeta) d\zeta &\geq \int_{\frac{\delta}{2}}^{\delta} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(0, \zeta) d\zeta \\ &\geq \frac{\delta^{1-\frac{\lambda_1}{a}}}{2} c e^{-\lambda_1 \delta r_{\infty}} > 0. \end{aligned}$$

And since $r(\eta)$ is bounded on $(0, \delta)$, $\lim_{\eta \downarrow 0} (r(\eta)\eta^{\frac{\lambda_1}{a}}) = 0$. Part 1 of Theorem 3 follows then from Lemma 10.

Now, let $\lambda_1 > a > 0$. In case B, $\bar{\rho}(0) > 0$ by Lemma 8. In case C, $\bar{\rho}(0) = 0$ because 0 is the right endpoint of an open interval that does not contain any points from the support of $(\mu_i)_{i \in S}$.

Since $\lambda_1 > a > 0$, there is a small $\alpha > 0$ such that

$$\frac{\lambda_1}{a} (1 - \alpha) > 1.$$

Let $\eta \in (0, \delta)$ so that $\eta < \eta^{\alpha} < \delta$. Then,

$$(17) \quad \rho_1(\eta) = \left(\frac{\eta^{\frac{\lambda_1}{a}-1}}{a} + r(\eta)\eta^{\frac{\lambda_1}{a}} \right) \int_{\eta}^{\eta^{\alpha}} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta$$

$$(18) \quad + \left(\frac{\eta^{\frac{\lambda_1}{a}-1}}{a} + r(\eta)\eta^{\frac{\lambda_1}{a}} \right) \int_{\eta^{\alpha}}^{\vartheta} \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta.$$

The term in (18) is bounded from above by

$$\begin{aligned} & \left(\frac{\eta^{\frac{\lambda_1}{a}-1}}{a} + r_\infty \eta^{\frac{\lambda_1}{a}} \right) \left(\int_{\eta^\alpha}^\delta \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta + \int_\delta^\vartheta \zeta^{-\frac{\lambda_1}{a}} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta \right) \\ & \leq \left(\frac{\eta^{\frac{\lambda_1}{a}-1}}{a} + r_\infty \eta^{\frac{\lambda_1}{a}} \right) (\eta^{-\frac{\lambda_1}{a}} e^{\lambda_1 \delta r_\infty} \|\bar{\rho}\|_1 + \delta^{-\frac{\lambda_1}{a}} e^{\lambda_1 \delta r_\infty} \|\bar{\rho}\|_1) \\ & = e^{\lambda_1 \delta r_\infty} \|\bar{\rho}\|_1 \left(\frac{\eta^{\frac{\lambda_1}{a}(1-\alpha)-1}}{a} + r_\infty \eta^{\frac{\lambda_1}{a}(1-\alpha)} + \frac{\delta^{-\frac{\lambda_1}{a}}}{a} \eta^{\frac{\lambda_1}{a}-1} + r_\infty \delta^{-\frac{\lambda_1}{a}} \eta^{\frac{\lambda_1}{a}} \right), \end{aligned}$$

which converges to 0 as η approaches 0 from the right.

Since $\eta^\alpha < \delta$, the function $\zeta \mapsto \bar{\rho}(\zeta)E(\eta, \zeta)$ is continuous on $[\eta, \eta^\alpha]$. By the mean-value theorem for integration, there exists $\zeta_\eta \in (\eta, \eta^\alpha)$ such that the term to the right of the equality sign in (17) equals

$$\begin{aligned} & \left(\frac{\eta^{\frac{\lambda_1}{a}-1}}{a} + r(\eta) \eta^{\frac{\lambda_1}{a}} \right) \int_\eta^{\eta^\alpha} \zeta^{-\frac{\lambda_1}{a}} d\zeta \bar{\rho}(\zeta_\eta) E(\eta, \zeta_\eta) \\ & = \left(\frac{1}{a} (1 - \eta^{(1-\alpha)(\frac{\lambda_1}{a}-1)}) + r(\eta) (\eta - \eta^{\alpha+(1-\alpha)\frac{\lambda_1}{a}}) \right) \frac{a}{\lambda_1 - a} \bar{\rho}(\zeta_\eta) E(\eta, \zeta_\eta). \end{aligned}$$

Since $\zeta_\eta \in (\eta, \eta^\alpha)$ for any η , it is clear that $\lim_{\eta \downarrow 0} \zeta_\eta = 0$. Continuity of $\bar{\rho}$ at $\eta = 0$ and integrability of $r(x)$ on $(0, \delta)$ then imply that

$$\lim_{\eta \downarrow 0} \left(\bar{\rho}(\zeta_\eta) E(\eta, \zeta_\eta) \frac{a}{\lambda_1 - a} \right) = a \frac{\bar{\rho}(0)}{\lambda_1 - a}.$$

Furthermore,

$$\lim_{\eta \downarrow 0} \left(\frac{1}{a} (1 - \eta^{(1-\alpha)(\frac{\lambda_1}{a}-1)}) \right) = \frac{1}{a}.$$

Finally, for small $\eta > 0$, we have $\eta > \eta^{\alpha+(1-\alpha)\frac{\lambda_1}{a}}$. It follows that

$$|r(\eta)| (\eta - \eta^{\alpha+(1-\alpha)\frac{\lambda_1}{a}}) \leq r_\infty (\eta - \eta^{\alpha+(1-\alpha)\frac{\lambda_1}{a}}),$$

which converges to 0 as η approaches 0 from the right. This completes the proof of part 2 of Theorem 3.

Next, assume that $\lambda_1 = a$. For $\eta \in (0, \delta)$,

$$\begin{aligned} \rho_1(\eta) & = \left(\frac{1}{a} + r(\eta) \eta \right) \int_\eta^\delta \zeta^{-1} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta \\ & + \left(\frac{1}{a} + r(\eta) \eta \right) \int_\delta^\vartheta \zeta^{-1} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta. \end{aligned} \tag{19}$$

By Lemma 9, the term in (19) is bounded on $(0, \delta)$. In case B, $c \leq \bar{\rho}(\eta) \leq \bar{\rho}_\infty$ for any $\eta \in [0, \delta]$. Therefore,

$$\begin{aligned} & -c e^{-\lambda_1 \delta r_\infty} \ln(\eta) + c e^{-\lambda_1 \delta r_\infty} \ln(\delta) \\ & \leq \int_\eta^\delta \zeta^{-1} \bar{\rho}(\zeta) E(\eta, \zeta) d\zeta \\ & \leq -\bar{\rho}_\infty e^{\lambda_1 \delta r_\infty} \ln(\eta) + \bar{\rho}_\infty e^{\lambda_1 \delta r_\infty} \ln(\delta) \end{aligned}$$

for $\eta \in (0, \delta)$. As

$$\lim_{\eta \downarrow 0} \left(\frac{1}{a} + r(\eta) \eta \right) = \frac{1}{a},$$

this establishes part 3 of Theorem 3 for case B. In case C, we only have $\bar{\rho}(\eta) \leq \bar{\rho}_\infty$, which is why we obtain a weaker statement. The proof of part 4 is similar to the proof of part 2 and we omit it. \square

7.2. Asymptotic analysis for analytic vector fields. In this subsection, we prove Theorem 2. The ensuing paragraph follows [BHK⁺11].

For any $i \in S$, we introduce the probability flux

$$\varphi_i(\eta) = \rho_i(\eta)u_i(\eta).$$

The vector of probability fluxes $(\varphi_1(\eta), \dots, \varphi_n(\eta))^T$ is denoted by $\varphi(\eta)$. As in Subsection 7.1, we let $\delta > 0$ be so small that the vector fields $(u_i)_{i>1}$ have no critical point in $[0, \delta]$ and u_1 has no critical point in $(0, \delta]$.

Since the invariant densities $(\rho_i)_{i \in S}$ are \mathcal{C}^1 on $(0, \delta)$, they satisfy the Fokker–Planck equations

$$(20) \quad \rho'_i(\eta)u_i(\eta) + \rho_i(\eta)u'_i(\eta) = -\lambda_i\rho_i(\eta) + \sum_{l \neq i} \lambda_{l,i}\rho_l(\eta), \quad i \in S,$$

on $(0, \delta)$, see [FGRC09]. Written in terms of the probability fluxes, (20) becomes

$$(21) \quad \varphi'_i(\eta) = -\frac{\lambda_i}{u_i(\eta)}\varphi_i(\eta) + \sum_{l \neq i} \frac{\lambda_{l,i}}{u_l(\eta)}\varphi_l(\eta), \quad i \in S.$$

In Appendix A, we show how Equation (16) can be derived directly from the Fokker–Planck equations if the invariant densities are \mathcal{C}^1 .

Our approach is to derive the asymptotically dominant term for the probability flux φ_1 , which will then immediately give the asymptotically dominant term for ρ_1 . We begin by showing that $\lim_{\eta \downarrow 0} \varphi_1(\eta) = 0$.

Lemma 11. *We have $\lim_{\eta \downarrow 0} \varphi_1(\eta) = 0$.*

PROOF: By Remark 6, the limit $\lim_{\eta \downarrow 0} \varphi_i(\eta)$ exists for any $i > 1$.

It is an easy corollary of (21) that

$$\sum_{i \in S} \varphi'_i(\eta) = 0$$

for any $\eta \in (0, \delta)$. Thus, the sum of all probability fluxes is equal to a constant k on this interval. Since

$$\varphi_1(\eta) = k - \sum_{i>1} \varphi_i(\eta)$$

for any $\eta \in (0, \delta)$, the limit $l := \lim_{\eta \downarrow 0} \varphi_1(\eta)$ exists as well. It remains to show that $l = 0$.

To obtain a contradiction, assume that $l \neq 0$. Then, there is no loss of generality in assuming that

$$|\varphi_1(\eta)| \geq \frac{|l|}{2}$$

for any $\eta \in (0, \delta)$. Since $u_1(\eta) = -a\eta + o(\eta)$ as η approaches 0 from the right, we may also assume that

$$\left| \frac{u_1(\eta)}{\eta} \right| \leq 2|a|, \quad \eta \in (0, \delta).$$

But this yields

$$\int_0^\delta \rho_1(\eta) d\eta = \int_0^\delta \frac{|\varphi_1(\eta)|}{|u_1(\eta)|} d\eta \geq \frac{|l|}{4|a|} \int_0^\delta \frac{d\eta}{\eta} = \infty,$$

which contradicts the fact that ρ_1 is integrable. \square

Corollary 4. *In case C from Section 3.2, $\lim_{\eta \downarrow 0} \varphi(\eta) = 0$.*

PROOF: In case C, the invariant densities $(\rho_i)_{i \in S}$ vanish to the left of 0. By Remark 6, the densities $(\rho_i)_{i > 1}$ are continuous at 0, which implies that $\lim_{\eta \downarrow 0} \rho_i(\eta) = 0$ for any $i > 1$. Hence, $\lim_{\eta \downarrow 0} \varphi_i(\eta) = 0$ for any $i > 1$, and $\lim_{\eta \downarrow 0} \varphi_1(\eta) = 0$ by Lemma 11. \square

We introduce the matrix of switching rates

$$\Lambda := \begin{pmatrix} -\lambda_1 & \lambda_{2,1} & \cdots & \lambda_{n,1} \\ \lambda_{1,2} & -\lambda_2 & \cdots & \lambda_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{1,n} & \lambda_{2,n} & \cdots & -\lambda_n \end{pmatrix},$$

and let $U(\eta)$ be the diagonal matrix with diagonal entries $\frac{1}{u_1(\eta)}, \dots, \frac{1}{u_n(\eta)}$.

For a fixed $\epsilon \in (0, \delta)$, we consider the initial-value problem

$$(22) \quad \begin{aligned} \phi'(\eta) &= \Lambda U(\eta) \phi(\eta), \\ \phi(\epsilon) &= \varphi(\epsilon), \end{aligned}$$

whose unique solution is $\varphi(\eta)$. Initial-value problem (22) can be written equivalently as

$$(23) \quad \begin{aligned} \phi'(\eta) &= \frac{1}{\eta} B(\eta) \phi(\eta), \\ \phi(\epsilon) &= \varphi(\epsilon). \end{aligned}$$

Here,

$$B(\eta) := \Lambda \tilde{U}(\eta),$$

where $\tilde{U}(\eta)$ is the diagonal matrix with diagonal entries $\frac{\eta}{u_1(\eta)}, \dots, \frac{\eta}{u_n(\eta)}$. Note that $B(\eta)$ is analytic at $\eta = 0$. This follows from the fact that the diagonal entries of $\tilde{U}(\eta)$ are analytic at $\eta = 0$, which is easily derived from analyticity of the vector fields. The linear system (23) then has a so-called regular singular point at $\eta = 0$ (see [Tay11, Section 3.11]).

Since $B(\eta)$ is analytic at $\eta = 0$, there exist a $\rho \in (0, \delta)$ and a sequence of matrices $(B_k)_{k \geq 0}$ such that

$$(24) \quad B(\eta) = \sum_{k=0}^{\infty} \eta^k B_k$$

for any $\eta \in (-\rho, \rho)$. There is no loss of generality in assuming that $\rho = \delta$.

Since $u_1(\eta) = -a\eta + O(\eta^2)$, and since $u_i(\eta) \neq 0$ for any $i > 1$, the matrix B_0 in (24) has the form

$$B_0 = \begin{pmatrix} \frac{\lambda_1}{a} & 0 & \cdots & 0 \\ -\frac{\lambda_{1,2}}{a} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\lambda_{1,n}}{a} & 0 & \cdots & 0 \end{pmatrix}.$$

It is easy to give a complete description of the eigenvalues and corresponding eigenspaces of B_0 .

Lemma 12. *The matrix B_0 has eigenvalues $\frac{\lambda_1}{a}$ and 0. The eigenspace corresponding to $\frac{\lambda_1}{a}$ is spanned by the vector $\lambda := (\lambda_1, -\lambda_{1,2}, -\lambda_{1,3}, \dots, -\lambda_{1,n})^T$. The eigenspace corresponding to 0 is the orthogonal complement to the span of $\{(1, 0, \dots, 0)^T\}$.*

We omit the proof of Lemma 12.

At this point, we need to distinguish between two cases. First, assume that $\frac{\lambda_1}{a}$ is not an integer. Such a condition is sometimes referred to as a nonresonance condition. The following statement is then a reformulation of [Tay11, Proposition 11.2].

Lemma 13. *There is a function*

$$(25) \quad V(\eta) = \mathbb{1} + \sum_{k=1}^{\infty} \eta^k V_k$$

that satisfies the normal equation

$$(26) \quad \eta V'(\eta) = B(\eta)V(\eta) - V(\eta)B_0, \quad \eta \in (0, \delta)$$

and for which

$$\varphi(\eta) = V(\eta) \exp\left(\ln\left(\frac{\eta}{\epsilon}\right) B_0\right) V(\epsilon)^{-1} \varphi(\epsilon), \quad \eta \in (0, \delta).$$

Now, we consider the resonance case, i.e we assume that $\frac{\lambda_1}{a}$ is a positive integer. In this case, we may not be able to construct a solution of the form (25) to (26). Instead, we consider the modified version

$$(27) \quad \eta V'(\eta) = B(\eta)V(\eta) - V(\eta)(B_0 + \eta^{\frac{\lambda_1}{a}} Y),$$

where Y is a matrix satisfying

$$(28) \quad B_0 Y = Y \left(B_0 + \frac{\lambda_1}{a} \mathbb{1} \right).$$

In this setting, we have the following reformulation of [Tay11, Proposition 11.5].

Lemma 14. *There exist a function $V(\eta)$ of the form (25) and a matrix Y satisfying (28) such that $V(\eta)$ satisfies (27) with Y and*

$$\varphi(\eta) = V(\eta) \exp\left(\ln\left(\frac{\eta}{\epsilon}\right) B_0\right) \exp\left(\ln\left(\frac{\eta}{\epsilon}\right) Y\right) V(\epsilon)^{-1} \varphi(\epsilon), \quad \eta \in (0, \delta).$$

PROOF OF THEOREM 2: Comparing Theorems 2 and 3, we see that we only need to show part 2 for case C and part 3 for both cases.

Let $\nu \in \mathbb{R}$ and let $\tilde{y} \in \mathbb{R}^n$ with first component equal to 0 such that

$$V(\epsilon)^{-1}\varphi(\epsilon) = \nu\lambda + \tilde{y},$$

where λ was defined in Lemma 12.

In the nonresonance case, Lemma 12 implies that

$$\begin{aligned} \exp\left(\ln\left(\frac{\eta}{\epsilon}\right)B_0\right)V(\epsilon)^{-1}\varphi(\epsilon) &= \sum_{k=0}^{\infty} \frac{1}{k!} \left(\ln\left(\frac{\eta}{\epsilon}\right)\right)^k (\nu B_0^k \lambda + B_0^k \tilde{y}) \\ &= \tilde{y} + \nu\lambda + \sum_{k=1}^{\infty} \frac{1}{k!} \left(\ln\left(\frac{\eta}{\epsilon}\right)\right)^k \nu \left(\frac{\lambda_1}{a}\right)^k \lambda \\ &= \tilde{y} + \nu \exp\left(\frac{\lambda_1}{a} \ln\left(\frac{\eta}{\epsilon}\right)\right) \lambda \\ (29) \quad &= \tilde{y} + \nu \epsilon^{-\frac{\lambda_1}{a}} \eta^{\frac{\lambda_1}{a}} \lambda, \end{aligned}$$

so

$$(30) \quad \varphi(\eta) = \left(\mathbb{1} + \sum_{k=1}^{\infty} \eta^k V_k\right) (\tilde{y} + \nu \epsilon^{-\frac{\lambda_1}{a}} \eta^{\frac{\lambda_1}{a}} \lambda), \quad \eta \in (0, \delta)$$

by Lemma 13. From (30), we infer that

$$\tilde{y} = \lim_{\eta \downarrow 0} \varphi(\eta).$$

In case C, Corollary 4 implies that $\tilde{y} = 0$. If ν was equal to 0, it would then follow that $\varphi \equiv 0$ on $(0, \delta)$. This is impossible in light of Lemma 8. As a result,

$$\varphi(\eta) = \nu \epsilon^{-\frac{\lambda_1}{a}} \eta^{\frac{\lambda_1}{a}} \lambda + o(\eta^{\frac{\lambda_1}{a}})$$

as η approaches 0 from the right. This establishes part 2 of Theorem 2 for case C and under the assumption that $\frac{\lambda_1}{a}$ is not an integer.

In the resonance case, Proposition 11.6 in [Tay11] implies that $Y^2 = 0$, that $Y\lambda = 0$ and that $Y\tilde{y}$ is an eigenvector of B_0 corresponding to the eigenvalue $\frac{\lambda_1}{a}$. Together with Lemma 14, this yields

$$(31) \quad \begin{aligned} \varphi(\eta) &= V(\eta) \exp\left(\ln\left(\frac{\eta}{\epsilon}\right)B_0\right) (\nu\lambda + \tilde{y} + \ln\left(\frac{\eta}{\epsilon}\right)Y(\nu\lambda + \tilde{y})) \\ &= V(\eta) \left(\exp\left(\ln\left(\frac{\eta}{\epsilon}\right)B_0\right) (\nu\lambda + \tilde{y}) \right. \end{aligned}$$

$$(32) \quad \left. + \ln\left(\frac{\eta}{\epsilon}\right) \left(Y\tilde{y} + \sum_{k=1}^{\infty} \frac{1}{k!} \left(\ln\left(\frac{\eta}{\epsilon}\right)\right)^k \left(\frac{\lambda_1}{a}\right)^k Y\tilde{y}\right)\right).$$

Using (29) and (25), the term in (31) and (32) becomes

$$(33) \quad \left(\mathbb{1} + \sum_{k=1}^{\infty} \eta^k V_k\right) (\tilde{y} + \epsilon^{-\frac{\lambda_1}{a}} \eta^{\frac{\lambda_1}{a}} (\nu\lambda - \ln(\epsilon)Y\tilde{y}) + \epsilon^{-\frac{\lambda_1}{a}} \eta^{\frac{\lambda_1}{a}} \ln(\eta)Y\tilde{y}).$$

Let us first consider the situation where $\frac{\lambda_1}{a} > 1$. In case C, $\tilde{y} = 0$ and we obtain

$$\varphi(\eta) = \nu \epsilon^{-\frac{\lambda_1}{a}} \eta^{\frac{\lambda_1}{a}} \lambda + o(\eta^{\frac{\lambda_1}{a}})$$

as η approaches 0 from the right. Since $\nu \neq 0$ by Lemma 8, we have established part 2 of Theorem 2 for case C under the assumption that $\frac{\lambda_1}{a}$ is an integer larger than 1.

Now, suppose that $\frac{\lambda_1}{a} = 1$. In case C, Representation (33) of $\varphi(\eta)$ implies that

$$\varphi(\eta) = \nu \epsilon^{-1} \eta \lambda + o(\eta),$$

and part 3 of Theorem 2 follows for case C.

In case B, (33) yields

$$\varphi(\eta) = \tilde{y} + \epsilon^{-1} \eta \ln(\eta) Y \tilde{y} + o(\eta \ln(\eta)).$$

Since $Y \tilde{y}$ is an eigenvector of B_0 with corresponding eigenvalue $\frac{\lambda_1}{a}$, Lemma 12 implies that the first component of $Y \tilde{y}$ is nonzero. This yields part 3 of Theorem 2 for case B. \square

8. PROOF OF LEMMAS 1 AND 2

For $t \geq 0$, as defined in (1) let $\mu \mathbf{P}^t$ denote the distribution of $(X, A)_t$ starting from the initial distribution μ . Since μ is invariant under $(\mathbf{P}^t)_{t \geq 0}$,

$$(34) \quad \mu_i(\cdot) = \int_{T^{(1)}}^{T^{(2)}} \pi(t) \mu \mathbf{P}^t(\times \{i\}) dt$$

for any $T^{(1)} < T^{(2)}$ in $[0, \infty]$ and for any probability density $\pi(t)$ on $(T^{(1)}, T^{(2)})$.

We will expand the expression on the right with respect to the sequences of driving vector fields and will ultimately see how ρ_i gets transformed through the action of the Markov semigroup and through time-averaging.

The following formula is the key to Lemmas 1 and 2.

Lemma 15. *Let $E \subset \mathbb{R}$ be a Borel set and let $i \in S$. For any $T^{(1)} < T^{(2)}$ in $[0, \infty]$ and for any probability density $\pi(t)$ on $(T^{(1)}, T^{(2)})$,*

$$\begin{aligned} \mu_i(E) &= \int_E \left(\int_{T^{(1)}}^{T^{(2)}} \pi(t) e^{-\lambda_i t} \Phi_i^t \# \rho_i(\eta) dt \right. \\ &\quad \left. + \sum_{j \neq i} \lambda_{j,i} \left(\int_0^{T^{(1)}} e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt + \int_{T^{(1)}}^{T^{(2)}} c(t) e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt \right) \right) d\eta, \end{aligned}$$

where $c(t) := \int_0^{T^{(2)}-t} \pi(s+t) ds$.

Given this representation for μ_i , we first show Lemma 1 and then Lemma 2. Finally, we prove the representation itself.

8.1. Proof of Lemma 1. When we set $T^{(1)} = 0$, $T^{(2)} = T$ and $\pi(t) = \frac{1}{T}$, the identity in Lemma 15 becomes

$$\begin{aligned} \mu_i(E) &= \int_E \left(\frac{1}{T} \int_0^T e^{-\lambda_i t} \Phi_i^t \# \rho_i(\eta) dt \right. \\ &\quad \left. + \sum_{j \neq i} \lambda_{j,i} \frac{1}{T} \int_0^T (T-t) e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt \right) d\eta. \end{aligned}$$

This implies Lemma 1.

8.2. Proof of Lemma 2. When we set $T^{(1)} = T$ for some time $T > 0$, $T^{(2)} = \infty$ and $\pi(t) = e^{T-t}$, the identity in Lemma 15 becomes

$$\begin{aligned} \mu_i(E) &= \int_E \left(\int_T^\infty e^{T-t} e^{-\lambda_i t} \Phi_i^t \# \rho_i(\eta) dt \right. \\ &\quad \left. + \sum_{j \neq i} \lambda_{j,i} \left(\int_0^T e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt + \int_T^\infty e^{T-t} e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt \right) \right) d\eta. \end{aligned}$$

Since μ is a probability measure,

$$\begin{aligned} \int_E \int_T^\infty e^{T-t} e^{-\lambda_i t} \Phi_i^t \# \rho_i(\eta) dt d\eta &= \int_T^\infty e^{T-t} e^{-\lambda_i t} \mu_i((\Phi_i^t)^{-1}(E)) dt \\ &\leq e^{-\lambda_i T} \int_T^\infty e^{T-t} dt = e^{-\lambda_i T}, \end{aligned}$$

where one should observe that the set $(\Phi_i^t)^{-1}(E)$ is well-defined even if $\Phi_i^{-t}(\eta)$ is undefined for some $\eta \in E$. Similarly,

$$\sum_{j \neq i} \lambda_{j,i} \int_E \int_T^\infty e^{T-t} e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt d\eta \leq \sum_{j \neq i} \lambda_{j,i} e^{-\lambda_i T}.$$

Letting T go to infinity, we obtain that

$$\mu_i(E) = \int_E \sum_{j \neq i} \lambda_{j,i} \int_0^\infty e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt d\eta,$$

and Lemma 2 follows.

8.3. Proof of Lemma 15. Fix an $i \in S$. We introduce some notation. For any $t > 0$ and for any index sequence \mathbf{i} with terminal index i , let $C_{\mathbf{i}}^t$ denote the event that the driving vector fields up to time t appear in the order given by \mathbf{i} . For any index sequence $\mathbf{i} = (i_1, \dots, i_{m-1}, i)$ of length $m \geq 2$, let $\mathbf{P}_{\mathbf{i}}$ be the probability that the first m driving vector fields appear in the order given by \mathbf{i} , conditioned on u_{i_1} being the first driving vector field. For $T > 0$ and $m \in \mathbb{N}$, we define the simplex $\Delta_{T,m}$ as the interior of the convex hull of the vectors $0, Te_1, \dots, Te_m$ in \mathbb{R}^m . For any vector v with m components, no matter whether v is a vector of indices, switching times or switching rates, let $v^{(m-1)}$ denote the projection of v onto its first $(m-1)$ coordinates. Moreover, let $\|v\|_1$ be the sum of the coordinates of v and let $\langle \cdot, \cdot \rangle$ denote the Euclidean inner product on the space that fits the context (usually \mathbb{R}^{m-1} or \mathbb{R}^m).

Lemma 16. *For any $T^{(1)} < T^{(2)}$ in $[0, \infty]$ and for any function $\pi(t)$ that is non-negative and integrable on $(T^{(1)}, T^{(2)})$,*

$$\begin{aligned} \int_{T^{(1)}}^{T^{(2)}} \pi(t) \mathbf{P}_{\xi,i}(C_{(i)}^t) \int_{\mathbb{R}} \mathbf{P}_{\xi,i}(X_t \in E | C_{(i)}^t) \mu_i(d\xi) dt \\ = \int_E \int_{T^{(1)}}^{T^{(2)}} \pi(t) e^{-\lambda_i t} \Phi_i^t \# \rho_i(\eta) dt d\eta. \end{aligned}$$

PROOF: This is immediate. \square

Lemma 17. *For any index sequence $\mathbf{i} = (i_1, \dots, i_{m-1}, i)$ of length $m \geq 2$, for any $T^{(1)} < T^{(2)}$ in $[0, \infty]$ and for any function $\pi(t)$ that is nonnegative and integrable on $(T^{(1)}, T^{(2)})$,*

$$\begin{aligned} & \int_{T^{(1)}}^{T^{(2)}} \pi(t) \mathbb{P}_{\xi, i_1}(C_{\mathbf{i}}^t) \int_{\mathbb{R}} \mathbb{P}_{\xi, i_1}(X_t \in E | C_{\mathbf{i}}^t) \mu_{i_1}(d\xi) dt \\ &= \mathbb{P}_{\mathbf{i}} \prod_{l=1}^{m-1} \lambda_{i_l} \int_{\Delta_{T^{(2)}, m} \setminus \Delta_{T^{(1)}, m}} \pi(\|\mathbf{s}\|_1) e^{-\langle \lambda^{(m-1)}, \mathbf{s}^{(m-1)} \rangle} e^{-\lambda_i s_m} \mu_{i_1}((\Phi_{\mathbf{i}}^{\mathbf{s}})^{-1}(E)) ds, \end{aligned}$$

where $\lambda^{(m-1)} = (\lambda_{i_1}, \dots, \lambda_{i_{m-1}})^T$.

PROOF: Fix an index sequence $\mathbf{i} = (i_1, \dots, i_{m-1}, i)$ of length m , $T^{(1)} < T^{(2)} \in [0, \infty]$ and a nonnegative integrable function π on $(T^{(1)}, T^{(2)})$. Let T_1, \dots, T_m be independent, exponentially distributed random variables such that T_l has parameter λ_{i_l} for $1 \leq l \leq m-1$ and T_m has parameter λ_i . For any $t \geq T^{(1)}$,

$$\begin{aligned} (35) \quad & \int_{\mathbb{R}} \mathbb{P}_{\xi, i_1}(X_t \in E | C_{\mathbf{i}}^t) \mu_{i_1}(d\xi) \\ &= \frac{1}{\mathbb{P}(R_{\mathbf{i}}^t)} \int_{\mathbb{R}} \mathbb{P}(\Phi_{\mathbf{i}}^{(T_1, \dots, T_{m-1}, t - \sum_{l=1}^{m-1} T_l)}(\xi) \in E, R_{\mathbf{i}}^t) \mu_{i_1}(d\xi), \end{aligned}$$

where

$$R_{\mathbf{i}}^t := \left\{ \sum_{l=1}^{m-1} T_l < t \leq \sum_{l=1}^m T_l \right\}.$$

As a notational shorthand, we introduce the functions

$$f_{t, \mathbf{i}}^{\xi}: \mathbb{R}^{m-1} \rightarrow \mathbb{R}, (s_1, \dots, s_{m-1}) \mapsto \Phi_{\mathbf{i}}^{(s_1, \dots, s_{m-1}, t - \sum_{l=1}^{m-1} s_l)}(\xi).$$

Then,

$$\begin{aligned} & \mathbb{P}(\Phi_{\mathbf{i}}^{(T_1, \dots, T_{m-1}, t - \sum_{l=1}^{m-1} T_l)}(\xi) \in E; R_{\mathbf{i}}^t) \\ &= \int_{\Delta_{t, m-1}} \mathbb{1}_{\{f_{t, \mathbf{i}}^{\xi}(\mathbf{s}) \in E\}}(\mathbf{s}) \prod_{l=1}^{m-1} \lambda_{i_l} e^{-\lambda_{i_l} s_l} e^{-\lambda_i (t - \|\mathbf{s}\|_1)} ds, \end{aligned}$$

which implies that (35) can be written as

$$(36) \quad \frac{1}{\mathbb{P}(R_{\mathbf{i}}^t)} \int_{\mathbb{R}} \int_{\Delta_{t, m-1}} \mathbb{1}_{\{f_{t, \mathbf{i}}^{\xi}(\mathbf{s}) \in E\}}(\mathbf{s}) \prod_{l=1}^{m-1} \lambda_{i_l} e^{-\lambda_{i_l} s_l} e^{-\lambda_i (t - \|\mathbf{s}\|_1)} ds \mu_{i_1}(d\xi).$$

Interchanging the order of integration, the righthand side of (36) becomes

$$\frac{1}{\mathbb{P}(R_{\mathbf{i}}^t)} \int_{\Delta_{t, m-1}} \prod_{l=1}^{m-1} \lambda_{i_l} e^{-\lambda_{i_l} s_l - \lambda_i (t - \|\mathbf{s}\|_1)} \mu_{i_1}((f_{t, \mathbf{i}}(\mathbf{s}))^{-1}(E)) ds.$$

We have thus shown that

$$\begin{aligned} & \int_{T^{(1)}}^{T^{(2)}} \pi(t) \mathbb{P}_{\xi, i_1}(C_{\mathbf{i}}^t) \int_{\mathbb{R}} \mathbb{P}_{\xi, i_1}(X_t \in E | C_{\mathbf{i}}^t) \mu_{i_1}(d\xi) dt = \int_{T^{(1)}}^{T^{(2)}} \pi(t) \frac{\mathbb{P}_{\xi, i_1}(C_{\mathbf{i}}^t)}{\mathbb{P}(R_{\mathbf{i}}^t)} \\ & \quad \cdot \int_{\Delta_{t, m-1}} \prod_{l=1}^{m-1} \lambda_{i_l} e^{-\lambda_{i_l} s_l - \lambda_i (t - \|\mathbf{s}\|_1)} \mu_{i_1}((f_{t, \mathbf{i}}(\mathbf{s}))^{-1}(E)) ds dt. \end{aligned}$$

The term $\frac{\mathbf{P}_{\xi, i_1}(C_{\mathbf{i}}^t)}{\mathbf{P}(R_{\mathbf{i}}^t)}$ gives the probability that the first m driving vector fields appear according to index sequence \mathbf{i} , conditioned on u_{i_1} being the first driving vector field. It is therefore equal to $\mathbf{P}_{\mathbf{i}}$. Interchanging the order of integration and substituting $s_m = t - \|\mathbf{s}\|_1$, Lemma 17 follows. \square

Lemma 18. *For any index sequence $\mathbf{i} = (i_1, \dots, i_{m-1}, i)$ of length $m \geq 2$, for any $T^{(1)} < T^{(2)}$ in $[0, \infty]$ and for any function $\pi(t)$ that is nonnegative and integrable on $(T^{(1)}, T^{(2)})$,*

$$\begin{aligned} \int_{T^{(1)}}^{T^{(2)}} \pi(t) \mathbf{P}_{\xi, i_1}(C_{\mathbf{i}}^t) \int_{\mathbb{R}} \mathbf{P}_{\xi, i_1}(X_t \in E | C_{\mathbf{i}}^t) \mu_{i_1}(d\xi) dt = \\ \int_{\Delta_{T^{(2)}, 2} \setminus \Delta_{T^{(1)}, 2}} \lambda_{i_{m-1}, i} e^{-\lambda_i t} \pi(s+t) \mathbf{P}_{\xi, i_1}(C_{\mathbf{i}(m-1)}^s) \\ \cdot \int_{\mathbb{R}} \mathbf{P}_{\xi, i_1}(X_s \in (\Phi_{\mathbf{i}}^t)^{-1}(E) | C_{\mathbf{i}(m-1)}^s) \mu_{i_1}(d\xi) d(s, t). \end{aligned}$$

PROOF: For notational compactness momentarily we introduce the notation $\tilde{\Delta}_i(m, t) := \Delta_{T^{(i)}-t, m-1}$. Then by Tonelli's theorem, the term to the right of the equality sign in Lemma 17 can be written as

$$\begin{aligned} \int_0^{T^{(1)}} \mathbf{P}_{\mathbf{i}} \prod_{l=1}^{m-1} \lambda_{i_l} \int_{(\tilde{\Delta}_2 \setminus \tilde{\Delta}_1)(m, t)} \pi(\|\mathbf{s}\|_1 + t) e^{-\langle \lambda^{(m-1)}, \mathbf{s} \rangle - \lambda_i t} \mu_{i_1}((\Phi_{\mathbf{i}}^{(\mathbf{s}, t)})^{-1}(E)) d\mathbf{s} dt \\ + \int_{T^{(1)}}^{T^{(2)}} \mathbf{P}_{\mathbf{i}} \prod_{l=1}^{m-1} \lambda_{i_l} \int_{\tilde{\Delta}_2(m, t)} \pi(\|\mathbf{s}\|_1 + t) e^{-\langle \lambda^{(m-1)}, \mathbf{s} \rangle - \lambda_i t} \\ \cdot \mu_{i_1}((\Phi_{\mathbf{i}}^{(\mathbf{s}, t)})^{-1}(E)) d\mathbf{s} dt \\ = \int_0^{T^{(1)}} \lambda_{i_{m-1}, i} e^{-\lambda_i t} \mathbf{P}_{\mathbf{i}(m-1)} \prod_{l=1}^{m-2} \lambda_{i_l} \int_{(\tilde{\Delta}_2 \setminus \tilde{\Delta}_1)(m, t)} \pi_t(\|\mathbf{s}\|_1) e^{-\langle \lambda^{(m-1)}, \mathbf{s} \rangle} \\ \cdot \mu_{i_1}((\Phi_{\mathbf{i}(m-1)}^{\mathbf{s}})^{-1}((\Phi_{\mathbf{i}}^t)^{-1}(E))) d\mathbf{s} dt \\ + \int_{T^{(1)}}^{T^{(2)}} \lambda_{i_{m-1}, i} e^{-\lambda_i t} \mathbf{P}_{\mathbf{i}(m-1)} \prod_{l=1}^{m-2} \lambda_{i_l} \int_{\tilde{\Delta}_2(m, t)} \pi_t(\|\mathbf{s}\|_1) e^{-\langle \lambda^{(m-1)}, \mathbf{s} \rangle} \\ \cdot \mu_{i_1}((\Phi_{\mathbf{i}(m-1)}^{\mathbf{s}})^{-1}((\Phi_{\mathbf{i}}^t)^{-1}(E))) d\mathbf{s} dt, \end{aligned}$$

where the function $\pi_t(s) := \pi(s+t)$ is nonnegative and integrable on $(T^{(1)}-t, T^{(2)}-t)$ if $t < T^{(1)}$, and is nonnegative and integrable on $(0, T^{(2)}-t)$ if $t > T^{(1)}$.

By another application of Lemma 17 (if $m > 2$) or of Lemma 16 (if $m = 2$), the previous term becomes

$$\begin{aligned}
& \int_0^{T^{(1)}} \lambda_{i_{m-1},i} e^{-\lambda_i t} \int_{T^{(1)}-t}^{T^{(2)}-t} \pi_t(s) \mathbb{P}_{\xi, i_1}(C_{\mathbf{i}^{(m-1)}}^s) \\
& \quad \cdot \int_{\mathbb{R}} \mathbb{P}_{\xi, i_1}(X_s \in (\Phi_i^t)^{-1}(E) | C_{\mathbf{i}^{(m-1)}}^s) \mu_{i_1}(d\xi) ds dt \\
& + \int_{T^{(1)}}^{T^{(2)}} \lambda_{i_{m-1},i} e^{-\lambda_i t} \int_0^{T^{(2)}-t} \pi_t(s) \mathbb{P}_{\xi, i_1}(C_{\mathbf{i}^{(m-1)}}^s) \\
& \quad \cdot \int_{\mathbb{R}} \mathbb{P}_{\xi, i_1}(X_s \in (\Phi_i^t)^{-1}(E) | C_{\mathbf{i}^{(m-1)}}^s) \mu_{i_1}(d\xi) ds dt \\
& = \int_{\Delta_{T^{(2)}, 2} \setminus \Delta_{T^{(1)}, 2}} \lambda_{i_{m-1},i} e^{-\lambda_i t} \pi(s+t) \mathbb{P}_{\xi, i_1}(C_{\mathbf{i}^{(m-1)}}^s) \\
& \quad \cdot \int_{\mathbb{R}} \mathbb{P}_{\xi, i_1}(X_s \in (\Phi_i^t)^{-1}(E) | C_{\mathbf{i}^{(m-1)}}^s) \mu_{i_1}(d\xi) d(s, t).
\end{aligned}$$

□

PROOF OF LEMMA 15: Fix a Borel set E , $T^{(1)} < T^{(2)} \in [0, \infty]$ and a probability density π on $(T^{(1)}, T^{(2)})$. Expanding the term to the right of the equality sign in (34) by conditioning on the sequences of driving vector fields, we obtain

$$\begin{aligned}
\mu_i(E) &= \int_{T^{(1)}}^{T^{(2)}} \pi(t) \mathbb{P}_{\xi, i}(C_{(i)}^t) \int_{\mathbb{R}} \mathbb{P}_{\xi, i}(X_t \in E | C_{(i)}^t) \mu_i(d\xi) dt \\
(37) \quad &+ \sum_{\mathbf{i}: |\mathbf{i}| \geq 2} \int_{T^{(1)}}^{T^{(2)}} \pi(t) \mathbb{P}_{\xi, i_1}(C_{\mathbf{i}}^t) \int_{\mathbb{R}} \mathbb{P}_{\xi, i_1}(X_t \in E | C_{\mathbf{i}}^t) \mu_{i_1}(d\xi) dt,
\end{aligned}$$

where $\sum_{\mathbf{i}: |\mathbf{i}| \geq 2}$ extends over all index sequences $\mathbf{i} = (i_1, \dots, i_{m-1}, i)$ with terminal index i and length $m \geq 2$.

By Lemma 16, it is enough to show that the term in (37) equals

$$\sum_{j \neq i} \lambda_{j,i} \int_E \left(\int_0^{T^{(1)}} e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt + \int_{T^{(1)}}^{T^{(2)}} c(t) e^{-\lambda_i t} \Phi_i^t \# \rho_j(\eta) dt \right) d\eta,$$

where $c(t)$ is defined as in Lemma 15. For any $m \geq 2$, let $\sum_{\mathbf{i}: |\mathbf{i}|=m}$ be the sum over all index sequences of length m with terminal index i . For any $j \in S$, let $\sum_{\mathbf{i}}^{(j)}$ be the sum over all index sequences \mathbf{i} with terminal index j . By Lemma 18, the term

in (37) can be written as

$$\begin{aligned}
(38) \quad & \sum_{m=2}^{\infty} \sum_{\mathbf{i}:|\mathbf{i}|=m} \int_{\Delta_{T^{(2)},2} \setminus \Delta_{T^{(1)},2}} \lambda_{i_{m-1},i} e^{-\lambda_i t} \pi(s+t) \mathbf{P}_{\xi,i_1}(C_{\mathbf{i}^{(m-1)}}^s) \\
& \cdot \int_{\mathbb{R}} \mathbf{P}_{\xi,i_1}(X_s \in (\Phi_i^t)^{-1}(E) | C_{\mathbf{i}^{(m-1)}}^s) \mu_{i_1}(d\xi) d(s,t) \\
& = \sum_{j \neq i} \int_{\Delta_{T^{(2)},2} \setminus \Delta_{T^{(1)},2}} \lambda_{j,i} e^{-\lambda_i t} \pi(s+t) \\
& \cdot \sum_{\mathbf{i}}^{(j)} \mathbf{P}_{\xi,i_1}(C_{\mathbf{i}}^s) \int_{\mathbb{R}} \mathbf{P}_{\xi,i_1}(X_s \in (\Phi_i^t)^{-1}(E) | C_{\mathbf{i}}^s) \mu_{i_1}(d\xi) d(s,t).
\end{aligned}$$

Moreover, for any fixed s ,

$$\sum_{\mathbf{i}}^{(j)} \mathbf{P}_{\xi,i_1}(C_{\mathbf{i}}^s) \int_{\mathbb{R}} \mathbf{P}_{\xi,i_1}(X_s \in (\Phi_i^t)^{-1}(E) | C_{\mathbf{i}}^s) \mu_{i_1}(d\xi) = \mu^{\mathbf{P}^s}((\Phi_i^t)^{-1}(E) \times \{j\}).$$

Since μ is invariant, $\mu^{\mathbf{P}^s}((\Phi_i^t)^{-1}(E) \times \{j\})$ equals $\mu_j((\Phi_i^t)^{-1}(E))$ and is thus independent of s .

As a result, the right side of (38) equals

$$\begin{aligned}
& \sum_{j \neq i} \int_{\Delta_{T^{(2)},2} \setminus \Delta_{T^{(1)},2}} \lambda_{j,i} e^{-\lambda_i t} \pi(s+t) \mu_j((\Phi_i^t)^{-1}(E)) d(s,t) \\
& = \sum_{j \neq i} \int_0^{T^{(1)}} \lambda_{j,i} e^{-\lambda_i t} \mu_j((\Phi_i^t)^{-1}(E)) \int_{T^{(1)}-t}^{T^{(2)}-t} \pi(s+t) ds dt \\
& \quad + \sum_{j \neq i} \int_{T^{(1)}}^{T^{(2)}} \lambda_{j,i} e^{-\lambda_i t} \mu_j((\Phi_i^t)^{-1}(E)) \int_0^{T^{(2)}-t} \pi(s+t) ds dt,
\end{aligned}$$

and Lemma 15 follows. \square

APPENDIX A. HOW EQUATION (16) RELATES TO THE FOKKER-PLANCK EQUATIONS

Equation (16) in the proof of Lemma 10 can also be derived from the Fokker-Planck equations, but in order to do this, one needs to assume that the invariant densities are \mathcal{C}^1 .

It is an immediate consequence of (20) that

$$\bar{\rho}(\zeta) = (\lambda_1 + u_1'(\zeta))\rho_1(\zeta) + u_1(\zeta)\rho_1'(\zeta),$$

see [BHK⁺11]. Hence, the term to the right of the equality sign in (16) equals

$$(39) \quad - \frac{1}{u_1(\eta)} \int_{\eta}^{\vartheta} (\lambda_1 + u_1'(\zeta))\rho_1(\zeta) \exp\left(\lambda_1 \int_{\eta}^{\zeta} \frac{dx}{u_1(x)}\right) d\zeta$$

$$(40) \quad - \frac{1}{u_1(\eta)} \int_{\eta}^{\vartheta} \rho_1'(\zeta)u_1(\zeta) \exp\left(\lambda_1 \int_{\eta}^{\zeta} \frac{dx}{u_1(x)}\right) d\zeta.$$

As

$$\lim_{\zeta \uparrow \vartheta} \left(\rho_1(\zeta) u_1(\zeta) \exp \left(\lambda_1 \int_{\eta}^{\zeta} \frac{dx}{u_1(x)} \right) \right) = 0$$

if u_1 is smooth and forward-complete, integration by parts implies that the term in (40) equals

$$(41) \quad \rho_1(\eta) + \frac{1}{u_1(\eta)} \int_{\eta}^{\vartheta} (\lambda_1 + u_1'(\zeta)) \rho_1(\zeta) \exp \left(\lambda_1 \int_{\eta}^{\zeta} \frac{dx}{u_1(x)} \right) d\zeta.$$

Since the second term in (41) cancels with the term in (39), we obtain (16).

REFERENCES

- [BH12] Yuri Bakhtin and Tobias Hurth. Invariant densities for dynamical systems with random switching. *Nonlinearity*, 25(10):2937–2952, 2012.
- [BHK⁺11] M. Balázs, G. Horváth, S. Kolumbán, P. Kovács, and M. Telek. Fluid level dependent markov fluid models with continuous zero transition. *Performance Evaluation*, 68(11):1149 – 1161, 2011. Special Issue: Performance 2011.
- [BLBMZ12a] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. Qualitative properties of certain piecewise deterministic Markov processes. Available at <http://arxiv.org/abs/1204.4143>, 2012.
- [BLBMZ12b] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. Quantitative ergodicity for some switched dynamical systems. *Electron. Commun. Probab.*, 17:no. 56, 14, 2012.
- [CFS82] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinai. *Ergodic theory*, volume 245 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskiĭ.
- [CH13] B. Cloez and M. Hairer. Exponential ergodicity for Markov processes with random switching. Available at <http://arxiv.org/abs/1303.6999>, 2013.
- [FGRC09] A. Faggionato, D. Gabrielli, and M. Ribezzi Crivellari. Non-equilibrium thermodynamics of piecewise deterministic Markov processes. *J. Stat. Phys.*, 137(2):259–304, 2009.
- [Hai06] Martin Hairer. Ergodic properties of Markov processes. Lectures given at the University of Warwick, <http://www.hairer.org/notes/Markov.pdf>, 2006.
- [HKNT98] G. Horton, V.G. Kulkarni, D.M. Nicol, and K.S. Trivedi. Fluid stochastic Petri nets: Theory, applications, and solution techniques. *European Journal of Operational Research*, 105:184–201, 1998.
- [Tay11] M. E. Taylor. *Introduction to Differential Equations*, volume 14 of *Pure and Applied Undergraduate Texts*. American Mathematical Society, Providence, RI, 2011.
- [YZ10] G. George Yin and Chao Zhu. *Hybrid switching diffusions*, volume 63 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2010. Properties and applications.

YURI BAKHTIN, SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA, USA

TOBIAS HURTH, SCHOOL OF MATHEMATICS, GEORGIA INSTITUTE OF TECHNOLOGY, ATLANTA GA, USA

JONATHAN C. MATTINGLY, MATHEMATICS DEPARTMENT, CENTER OF THEORETICAL AND MATHEMATICAL SCIENCE, DEPARTMENT OF STATISTICAL SCIENCE, AND CENTER OF NONLINEAR AND COMPLEX SYSTEMS, DUKE UNIVERSITY, DURHAM NC, USA