SINGULARITIES OF INVARIANT DENSITIES FOR RANDOM SWITCHING BETWEEN TWO LINEAR ODES IN 2D.

YURI BAKHTIN, TOBIAS HURTH, SEAN D. LAWLEY, JONATHAN C. MATTINGLY

Abstract. We consider a planar dynamical system generated by two stable linear vector fields with distinct fixed points and random switching between them. We characterize singularities of the invariant density in terms of the switching rates and contraction rates. We prove boundedness away from those singularities. We also discuss some motivating biological examples.

1. Introduction

This paper describes the formation of singularities and regularity properties in the stationary densities for the dynamics created by random switching between two linear ordinary differential equations (ODEs) in the two-dimensional plane. A full characterization of stationary density singularities for randomly switched ODEs in one dimension is provided in [3], yet singularity formation is poorly understood in higher dimensions, even at the level of motivational examples.

Here, we study a deceptively simple two-dimensional example in the hope that it will begin to illuminate a path forward. We have not sought generality; but rather, picked a simple switching system between two linear equations to explore how geometry of contraction and random switching interact to produce singularities in the longtime distribution of the system. Despite this apparent simplicity, the structure of the stationary density can be quite rich. Depending on the relationships between the switching rates and the contraction rates, the stationary density may be bounded, have isolated singularites, or have one-dimensional curves of singularities. Though we have studied a particular system, our methods are fairly general and hopefully can be extended to an interesting class of examples.

There has been a resurgence in the study of such switched ODE systems in recent years under the names *hybrid systems* [19], *piecewise deterministic Markov processes (PDMP)* [8, 16], and *random evolutions* [10]. Some of this renewed interest stems from applications in ecology and cellular biology [15, 6]. On the more theoretical side, it was shown in [1, 5, 4] that a combination of a condition of Hörmander type and an accessibility condition guarantees that an invariant distribution, if it exists, is unique and absolutely continuous with respect to the Lebesgue measure.

One could expect that, similarly to the well-known results for hypoelliptic diffusions based on pseudo-differential calculus or Malliavin calculus, the same Hörmander condition would guarantee C^{∞} smoothness of the invariant density if the driving vector fields are smooth and a hypoellipticity condition is met. As already alluded to, the picture is more involved and invariant densities of switching systems often have singularities. In [3], emergence of singularities of invariant densities for one-dimensional switching systems due to contraction near stable critical points was studied, and a classification of singularities was given. It was also shown that away from critical points of the driving vector fields, the invariant densities are C^∞.

In higher dimensions, the situation is even more complex generically. Some of the flows generated by the driving vector fields may exhibit long-term contraction with or without convergence to a stable critical point, e.g., there may be more sophisticated low-dimensional attractors. Density singularities created by some of the vector fields may be propagated in new directions by other vector fields. Additional complexity emerges due to the presence of manifolds of hypoellipticity points.

We started an exploration of higher dimensions in [2], where we considered a class of switching systems on the two-dimensional torus that is devoid of these obstacles (the contraction is subexponential and all points are elliptic). For this class, we showed that the invariant densities belong to C^∞ and that there are no singularities.

For generic switching systems, characterizing singularities of the invariant densities and proving smoothness away from those singularities still seems to be a hard problem. In the present paper, for the first time we consider switching systems with a whole line of points of hypoellipticity and contractive flows associated to the driving vector fields. In various regimes that we define in terms of the parameters of the model, i.e., contraction rates and switching rates, we describe points and lines of singularities of the invariant density and prove boundedness of the density away from those singularities.

Let us describe the system more precisely now. We consider the PDMP given by Poissonian random switching between the linear vector fields

$$u_i(x_1, x_2) = \begin{pmatrix} -\alpha & 0 \\ 0 & -\beta \end{pmatrix} \begin{pmatrix} x_1 - i \\ x_2 - i \end{pmatrix}, \quad i = 0, 1, \tag{1}$$

where $\alpha > \beta > 0$. Given a starting point $x \in \mathbb{R}^2$ and an initial vector field, say u_0, we follow the flow of u_0 for an exponential time. Then a switch occurs, meaning that the driving vector field u_0 is replaced with u_1. Starting from the point in $\mathbb{R}^2$ where the switch occurred, we flow along u_1 for another exponential time, then switch back to u_0, etc. We assume that the times between consecutive switches are independent. Switches from u_0 to u_1 happen at a constant rate $\lambda_0 > 0$, and switches from u_1 to u_0 happen at a constant rate $\lambda_1 > 0$. The resulting dynamics are strongly affected by the globally asymptotically stable equilibrium points $(0, 0)$ and $(1, 1)$ of the two vector fields: A typical switching trajectory obtained from intermittent switching between u_0 and u_1 enters in finite time the region Γ bounded by the trajectory of u_0 starting from $(1, 1)$ and the trajectory of u_1 starting from $(0, 0)$, and then remains in Γ for all future times (see Figure 1 below). Since the setting is essentially compact, the semigroup of the PDMP admits an invariant probability measure. As will be established rigorously in Proposition 1, the invariant probability measure is unique and has a density with respect to the product of Lebesgue measure on $\mathbb{R}^2$ and counting measure on $\{0, 1\}$. The goal of this article is to investigate the marginals ρ_0 and ρ_1 of the density, corresponding to the driving vector fields u_0 and u_1. In this introduction and throughout the paper, we use the term *invariant densities* for the marginals of the density associated with an absolutely continuous invariant probability measure.

The PDMP governed by u_0 and u_1 can be thought of as a two-dimensional version of one of the simplest possible switching systems on the real line: If we switch between $v_0(x) = -ax$ and $v_1(x) = a(1-x)$ for $a > 0$, the resulting switching trajectory is alternately attracted by 0 and 1. As in the more complex two-dimensional system in (1), this simple one-dimensional system gives rise to a unique and absolutely continuous invariant probability measure. Unlike the invariant densities in the 2D system, however, the invariant densities in the 1D system can be computed explicitly by solving the corresponding Kolmogorov forward equations, see e.g. [9]. They are densities of beta distributions:

$$\begin{aligned} \rho_0(x) =& c_0 x^{\frac{\lambda_0}{a}-1}(1-x)^{\frac{\lambda_1}{a}}, \\ \rho_1(x) =& c_1 x^{\frac{\lambda_0}{a}}(1-x)^{\frac{\lambda_1}{a}-1}, \end{aligned} \tag{2}$$

where c_0, c_1 are constants. In particular, ρ_0 and ρ_1 are smooth in the interior of $[0, 1]$, and develop singularities at the critical points 0 and 1 if the switching rates are small compared to the rate of contraction a. While it is possible to write down the Kolmogorov forward equations for the invariant densities of (1), we cannot find explicit solutions to the equations. Besides, it is a priori not clear whether the invariant densities of (1) are sufficiently regular to be classical solutions on some meaningful set, say in the interior of Γ. Notice, however, that the marginals of the invariant densities with respect to the coordinates x_1 and x_2 are explicitly given by the formulae in (2) for $a = \alpha$ and $a = \beta$. We conjecture that the invariant densities for (1) are C^∞ in the interior of the set Γ. At the boundary of Γ, singularities may form due to exponential contraction and thus accumulation of probabilistic mass near the critical points, and the subsequent propagation of mass along trajectories of u_0 and u_1.

We give two results on singularities of the invariant densities for slow switching. The first one describes the singularities near the attracting critical points of u_0 and u_1. The basic mechanism leading to these singularities is mass accumulation due to the fact that, under small to moderate switching rates, there are long time intervals during which the system is exposed to contraction towards $(0, 0)$ and $(1, 1)$. The second result holds only for small switching rates, and describes how a singularity at the critical point of u_1 is spread along the trajectory of u_0 passing through this critical point. These results on singularity formation are complemented by several boundedness results, a first step towards proving regularity of the invariant densities in the interior of Γ. For instance, we show that the invariant densities are bounded on any compact set contained in the interior of Γ, even if switches are rare.

There are two main difficulties in dealing with the switching system in (1). The main obstacle to showing smoothness of the invariant densities is arguably the exponential contraction in the vicinity of critical points. Another less obvious difficulty stems from the fact that the vector fields u_0 and u_1 are aligned with each other along the diagonal line segment connecting $(0, 0)$ and $(1, 1)$. This partial breakdown of transversality makes the smoothing effect of switches close to the diagonal less pronounced. On the other hand, switches close to the diagonal but far from the critical points at least do not spoil the densities, which seems to make them a technical nuisance rather than an essential obstacle to establishing smoothness.

The paper is organized as follows. In Section 2, we discuss two systems emerging in applications that can be reduced to (1). We state our results on singularities of

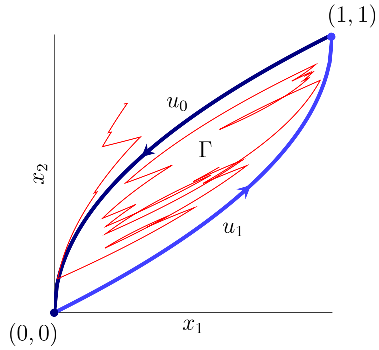

FIGURE 1. The support Γ of the invariant densities is the region bounded by the forward u_0 trajectory starting at $(1,1)$ and the forward u_1 trajectory starting at $(0,0)$. The red path shows a single stochastic realization of $(x_1(t), x_2(t))$.

the invariant density in Section 3. In Section 4, we prove existence and uniqueness of the invariant distribution, as well as a basic description of its support. Furthermore, we exhibit the line of hypoellipticity points, which is an obstacle to establishing boundedness of the invariant density. In Section 5, we recall some basic integral equations satisfied by the invariant density. In Section 6, we prove one of our main results (Theorem 1), which describes the singularities of the invariant density. In Section 7, we perform a change of variables in the integral equations from Section 5 that prepares the proof of our main boundedness result (Theorem 2). The latter is given in Section 8, where most of the technical work is carried out.

2. APPLICATIONS

Generically, our results concern any two-dimensional randomly switching ODE of the form

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathbf{b}_0 \mathbb{1}_{I_t=0} + \mathbf{b}_1 \mathbb{1}_{I_t=1}, \quad \mathbf{b}_0, \mathbf{b}_1 \in \mathbb{R}^2, \tag{3}$$

where $I_t \in \{0,1\}$ is a Markov jump process and $A \in \mathbb{R}^{2\times 2}$ has two distinct, negative eigenvalues. In particular, (3) reduces to (1) after the coordinate change

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = G\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} -\frac{1}{\alpha} & 0 \\ 0 & -\frac{1}{\beta} \end{pmatrix} G\mathbf{b}_0 \mathbb{1}_{I_t=0} + \left(\begin{pmatrix} -\frac{1}{\alpha} & 0 \\ 0 & -\frac{1}{\beta} \end{pmatrix} G\mathbf{b}_1 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \mathbb{1}_{I_t=1},$$

where G is an invertible (2×2) matrix such that

$$A = G^{-1}\begin{pmatrix} -\alpha & 0 \\ 0 & -\beta \end{pmatrix} G.$$

In addition to being one of the simplest nontrivial two-dimensional PDMP examples in which to study invariant densities, models of the form (3) arise naturally in diverse applications. We now give two such applications.

2.1. **Stochastic gene expression.** Much of the recent interest in PDMPs stems from their application to gene expression [12, 17, 6]. Models in this context typically begin with a continuous-time Markov chain on a discrete state space that tracks gene products (an integer number of mRNA and/or protein molecules) as well as some discrete (often binary) environmental state, such as whether a gene is active or inactive. Assuming that the number of gene products is large, one often approximates the amount of gene product by a continuous variable that evolves by a deterministic ODE between stochastic switches in the environmental state. That is, the stochasticity stemming from the finite number of gene products is averaged out, while the stochastic environmental state is retained.

To illustrate this concretely, we briefly describe the so-called "standard model" of gene expression [17]. Let $I_t \in \{0, 1\}$ be the state of a gene, with $I_t = 0$ ($I_t = 1$) corresponding to an active (inactive) gene, and suppose I_t leaves state $i \in \{0, 1\}$ at rate $\lambda_i > 0$. When the gene is active, it produces mRNA molecules at rate $\alpha > 0$. Each mRNA molecule degrades at rate $\delta > 0$ and produces a protein molecule at rate $\beta > 0$. Protein molecules degrade at rate $\gamma > 0$. Letting $X_t \in \{0, 1, 2, \dots\}$ and $Y_t \in \{0, 1, 2, \dots\}$ denote the respective mRNA and protein copy numbers, the Markov transitions are summarized by

$$I_t: \quad 0 \underset{\lambda_1}{\overset{\lambda_0}{\rightleftharpoons}} 1; \qquad X_t: \quad X \underset{\delta}{\overset{I_t\alpha}{\rightleftharpoons}} X+1, \qquad Y_t: \quad Y \underset{\gamma}{\overset{X_t\beta}{\rightleftharpoons}} Y+1. \tag{4}$$

This three-component Markov chain $(X_t, Y_t, I_t) \in \{0, 1, 2, \dots\}^2 \times \{0, 1\}$ and various simplifications have been very well studied using a variety of mathematical techniques [17, 6]. Indeed, depending on the parameter regime, this Markov chain has been reduced to an ODE, a PDMP, a stochastic differential equation (SDE) driven by white noise, an SDE driven by Lévy noise, and a Lévy-type process [11].

For our purposes, suppose that the characteristic number of mRNA and protein molecules is large,

$$X^* := \frac{\alpha}{\delta}\frac{\lambda_0}{\lambda_0 + \lambda_1} \gg 1, \quad Y^* := \frac{\beta}{\gamma}X^* \gg 1.$$

In this parameter regime, one can approximate the rescaled mRNA and protein concentrations, $x(t) := X_t/X^*$ and $y(t) := Y_t/Y^*$, by the two-dimensional PDMP [18],

$$\begin{aligned} \tfrac{d}{dt}x(t) &= \frac{\alpha}{X^*}I_t - \delta x(t), \\ \tfrac{d}{dt}y(t) &= \gamma(x(t) - y(t)), \end{aligned} \tag{5}$$

in which the only source of stochasticity remaining is I_t. Of course, (5) is of the form (3) if $\gamma \neq \delta$.

2.2. **PDEs with randomly switching boundary conditions.** While most of the interest in PDMPs has focused on switching ODEs, a number of biological applications have recently prompted the study of PDEs with randomly switching boundary conditions (for example, see [15, 14, 7, 13]). Perhaps the simplest such example is the one-dimensional diffusion equation,

$$\tfrac{\partial}{\partial t}c(x,t) = \tfrac{\partial^2}{\partial x^2}c(x,t), \quad x \in (0,1),$$

with an absorbing boundary condition at $x = 0$ and a randomly switching boundary condition at $x = 1$,

$$c(0,t) = 0, \quad c(1,t) = I_t,$$

where $I_t \in \{0,1\}$ is a continuous-time Markov jump process. Writing the solution in terms of the $L^2[0,1]$-orthonormal basis, $\{\sqrt{2}\sin(n\pi x)\}_{n=1}^{\infty}$,

$$c(x,t) = \sum_{n=1}^{\infty} c_n(t)\sqrt{2}\sin(n\pi x),$$

it follows that any pair of coefficients, say $c_k(t)$ and $c_m(t)$, satisfy the two-dimensional switching ODEs,

$$\begin{aligned} \tfrac{d}{dt}c_k(t) &= -\beta_k(c_k(t) - I_t b_k), \\ \tfrac{d}{dt}c_m(t) &= -\beta_m(c_m(t) - I_t b_m), \end{aligned} \tag{6}$$

where $\beta_n = n^2\pi^2$ and $b_n = (-1)^{n+1}\sqrt{2}/(n\pi)$. Of course, (6) is of the form (3) if $k \neq m$.

3. Problem setting and main results

We consider random switching between the linear vector fields u_0 and u_1 on $\mathbb{R}^2$, given by

$$u_i(x) = u_i(x_1,x_2) = \begin{pmatrix} -\alpha & 0 \\ 0 & -\beta \end{pmatrix}\begin{pmatrix} x_1 - i \\ x_2 - i \end{pmatrix}, \quad i = 0,1,$$

where $\alpha > \beta > 0$. The vector fields u_0 and u_1 have an attracting critical point at $(0,0)$ and $(1,1)$, respectively. For any $(x_1,x_2) \in \mathbb{R}^2$, the initial-value problem

$$\begin{pmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \end{pmatrix} = u_i(x_1(t),x_2(t)), \quad \begin{pmatrix} x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

has the unique solution

$$\Phi_i^t(x_1,x_2) = \begin{pmatrix} i + (x_1 - i)e^{-\alpha t} \\ i + (x_2 - i)e^{-\beta t} \end{pmatrix}, \quad t \in \mathbb{R}. \tag{7}$$

It is easy to see that

$$\Phi_1^t(x_1,x_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \Phi_0^t(1-x_1, 1-x_2), \quad x \in \mathbb{R}^2,\ t \in \mathbb{R}. \tag{8}$$

For notational convenience, we also define the inverse flows

$$\Psi_i^t(x) = (\Phi_i^t)^{-1}(x) = \Phi_i^{-t}(x), \quad i \in \{0,1\},\ t \in \mathbb{R},\ x \in \mathbb{R}^2.$$

As we will be switching intermittently between u_0 and u_1, it is also convenient to define the cumulative flows

$$\Phi_i^{(t_1,\ldots,t_n)} = \begin{cases} \Phi_i^{t_n} \circ \Phi_{1-i}^{t_{n-1}} \circ \Phi_i^{t_{n-2}} \circ \ldots \circ \Phi_{1-i}^{t_1}, & n \equiv 0 \mod 2, \\ \Phi_i^{t_n} \circ \Phi_{1-i}^{t_{n-1}} \circ \Phi_i^{t_{n-2}} \circ \ldots \circ \Phi_i^{t_1}, & n \equiv 1 \mod 2 \end{cases}$$

and

$$\Psi_i^{(t_1,\ldots,t_n)} = \left(\Phi_i^{(t_1,\ldots,t_n)}\right)^{-1}.$$

For $i \in \{0,1\}$, we call the set $\{\Phi_i^t(x) :\ t > 0\}$ the forward u_i trajectory starting at x and we call $\{\Phi_i^t(x) :\ t < 0\}$ the backward u_i trajectory starting at x. The set $\{\Phi_i^t(x) :\ t \in \mathbb{R}\}$ is simply called the u_i trajectory through x.

Let $I = (I_t)_{t\geq 0}$ be a continuous-time Markov chain on $\{0,1\}$ with jump rate λ_0 from 0 to 1 and λ_1 from 1 to 0. Then, we define a stochastic process $X = (X_t)_{t\geq 0}$ on $\mathbb{R}^2$ via

$$\tag{9} \tfrac{d}{dt} X_t = u_{I_t}(X_t).$$

The two-component process (X, I) is a Markov process on $\mathbb{R}^2 \times \{0,1\}$, whose Markov semigroup we denote by $(\mathsf{P}^t)_{t\geq 0}$. We call a probability measure μ on $\mathbb{R}^2 \times \{0,1\}$ an *invariant probability measure* of $(\mathsf{P}^t)_{t\geq 0}$ if $\mu = \mu \mathsf{P}^t$ for all $t \geq 0$.

The forward u_0 trajectory starting at $(1,1)$ and the forward u_1 trajectory starting at $(0,0)$ together with the critical points $(0,0)$ and $(1,1)$ mark the boundary of the set

$$\Gamma = \left\{ (x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_2 \leq 1,\ x_2^{\frac{\alpha}{\beta}} \leq x_1 \leq 1 - (1 - x_2)^{\frac{\alpha}{\beta}} \right\}.$$

We denote the interior of Γ by Γ°. Notice that Γ and Γ° are symmetric about the point $(\frac{1}{2}, \frac{1}{2})$, i.e. $(x_1, x_2) \in \Gamma$ $(\in \Gamma^\circ)$ if and only if $(1 - x_1, 1 - x_2) \in \Gamma$ $(\in \Gamma^\circ)$.

Proposition 1. *The Markov semigroup $(\mathsf{P}^t)_{t\geq 0}$ admits a unique invariant probability measure μ. It is absolutely continuous with respect to the product of Lebesgue measure on $\mathbb{R}^2$ and counting measure on $\{0,1\}$. Moreover, the marginals $\mu_i(\cdot) = \mu(\cdot \times \{i\})$, $i \in \{0,1\}$, have support Γ.*

Recall that the support of μ_i is the collection of all points $x \in \mathbb{R}^2$ such that $\mu_i(U) > 0$ for every neighborhood U of x. We prove Proposition 1 in Section 4. Since the marginal μ_i, $i \in \{0,1\}$, is absolutely continuous with respect to Lebesgue measure, it has a density $\rho_i \in L^1(\mathbb{R}^2)$, which we call an *invariant density*. Below, we state our results on boundedness as well as the occurrence of singularities for the invariant density ρ_0. Exploiting the symmetries of the switching system, one can easily formulate corresponding results for ρ_1. Whether and where singularities of ρ_0 occur depends critically on the switching rates λ_0 and λ_1. In some sense this is not surprising because small switching rates translate into few switches and thus an accumulation of probabilistic mass at the critical points $(0,0)$ and $(1,1)$. Interestingly, if both λ_0 and λ_1 are very small, the singularity created at the critical point of one of the vector fields is propagated along the forward trajectory of the other vector field that starts at the critical point. As L^1 functions, ρ_0 and ρ_1 are only defined up to a set of Lebesgue measure zero, so when we state, e.g., that ρ_0 is bounded on a set S, we mean that there is a representative of ρ_0 that is bounded on S. Proposition 1 implies that ρ_0 and ρ_1 vanish outside of Γ, which is why we can restrict ourselves to Γ° instead of considering all of $\mathbb{R}^2$.

For $i \in \{0,1\}$, let $\partial\Gamma_i$ denote the forward u_i trajectory starting at $(1-i, 1-i)$, and set

$$\Gamma_i = \Gamma \setminus \{(i,i)\}.$$

Observe that $\partial\Gamma_i$, $i \in \{0,1\}$, are the curves that make up the right and left part of the boundary of Γ, minus the critical points $(0,0)$ and $(1,1)$. The following theorem describes for which switching rates and in which regions singularities occur.

Theorem 1. *The following statements hold.*

1. *For $\lambda_0 < \alpha + \beta$, the invariant density ρ_0 is unbounded in every neighborhood of $(0,0)$.*
2. *For $\lambda_1 < \beta$ and $x \in \partial\Gamma_0$, ρ_0 is unbounded in every neighborhood of x. Since being unbounded in every neighborhood of a point is a closed condition, ρ_0 is also unbounded in every neighborhood of $(0,0)$ and $(1,1)$.*

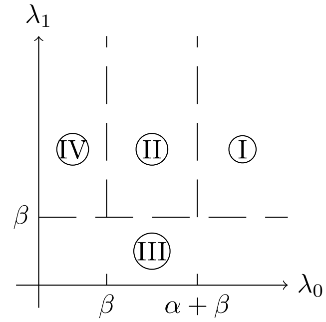

FIGURE 2. In the (λ_0, λ_1)-quadrant, the dashed lines delineate four regions whose labels indicate the following: I: ρ_0 is bounded everywhere on Γ°; II: ρ_0 is unbounded in every neighborhood of $(0,0)$, and bounded on every compact subset of Γ_0; III: for every $x \in \partial\Gamma_0$, ρ_0 is unbounded in every neighborhood of x; IV: classification is unclear. For switching rates on the dashed lines, the classification is unclear as well.

Theorem 2. *The following statements hold.*

(1) For $\lambda_0 > \alpha + \beta$ and $\lambda_1 > \beta$, the invariant density ρ_0 is bounded on Γ°.
(2) Let $\lambda_0, \lambda_1 > \beta$ and let $K \subset \Gamma_0$ be compact. Then, ρ_0 is bounded on K.
(3) Let $K \subset \Gamma$ be a compact set such that $K \cap \partial\Gamma_0 = \emptyset$. Then, ρ_0 is bounded on K for any switching rates $\lambda_0, \lambda_1 > 0$.

Remark 1. Theorems 1 and 2 do not address whether ρ_0 stays bounded along $\partial\Gamma_0$ if $\lambda_0 < \beta$ and $\lambda_1 > \beta$ (see Figure 2). Based on simulations (see the bottom right panel in Figure 3), we conjecture that ρ_0 is bounded in this case, i.e. we conjecture that Region II in Figure 2 contains all (λ_0, λ_1) such that $\lambda_0 < \alpha + \beta$ and $\lambda_1 > \beta$. The critical cases not covered by Theorems 1 and 2 (e.g., is ρ_0 bounded on Γ° if $\lambda_0 = \alpha + \beta$ and $\lambda_1 = \beta$?) are also open. These correspond to the dashed lines in Figure 2.

We prove Theorem 1 in Section 6. The proof of Theorem 2 is given in Section 8. Theorems 1 and 2 combined provide the following picture: For fast switching away from u_0 ($\lambda_0 > \alpha + \beta$) and for at least intermediate switching away from u_1 ($\lambda_1 > \beta$), the invariant density ρ_0 is globally bounded (see Region I in Figure 2). For intermediate switching away from u_0 ($\beta < \lambda_0 < \alpha+\beta$) and for at least intermediate switching away from u_1 ($\lambda_1 > \beta$), ρ_0 has a singularity at $(0,0)$, the critical point of u_0, but remains bounded away from $(0,0)$ (Region II in Figure 2). And in the regime of slow switching away from u_1 ($\lambda_1 < \beta$), ρ_0 has singularities along the entire left boundary curve of the support, including $(0,0)$ (Region III in Figure 2). This happens regardless of how quickly on average we switch away from u_0. Away from the left boundary curve, ρ_0 is always bounded. The mechanism leading to the blow-up of ρ_0 along $\partial\Gamma_0$ can be roughly described as follows: Due to exponential contraction of the flow of u_1, probabilistic mass accumulates at the sink $(1,1)$. This mass is subsequently propagated under the flow of u_0 and thus gives rise to singularities along the forward u_0 trajectory starting at $(1,1)$. These results are illustrated using stochastic simulations in Figure 3.

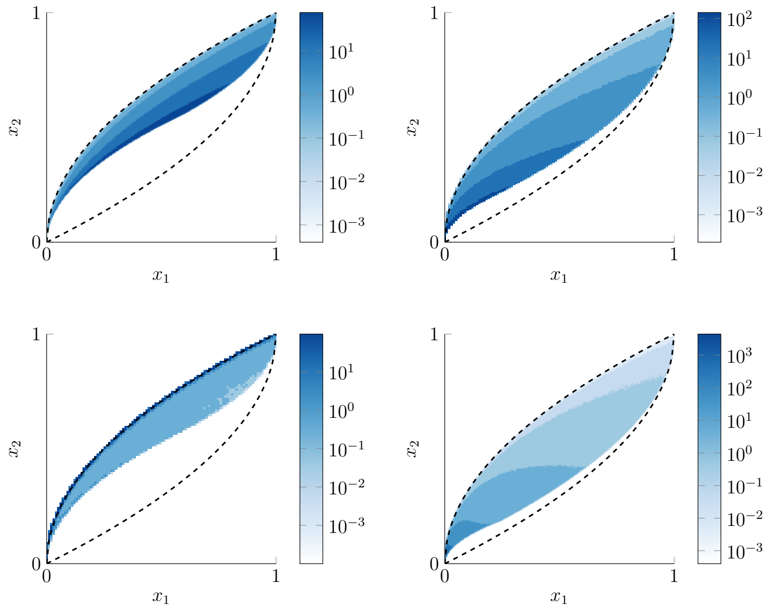

FIGURE 3. The four panels show the empirical probability density on a logarithmic scale obtained from stochastic simulations for different choices of the switching rates λ_0 and λ_1. The contraction rates are $\alpha = 2$ and $\beta = 1$ in all panels. The top left panel corresponds to Region I in Figure 2 with $\lambda_0 = 4$ and $\lambda_1 = 2$. The top right panel corresponds to Region II in Figure 2 with $\lambda_0 = \lambda_1 = 2$. The bottom left panel corresponds to Region III in Figure 2 with $\lambda_0 = 2$ and $\lambda_1 = 10^{-2}$. The bottom right panel corresponds to Region IV in Figure 2 with $\lambda_0 = 0.5$ and $\lambda_1 = 2$.

4. The support of the invariant measure

In this section, we prove Proposition 1. Given two points $x, y \in \mathbb{R}^2$, we say that x is *reachable* from y if there exist $i \in \{0,1\}$, $n \in \mathbb{N}$, and $(t_1, \dots, t_n) \in \mathbb{R}_+^n$ such that

$$x = \Phi_i^{(t_1,\dots,t_n)}(y).$$

For $y \in \mathbb{R}^2$, let $L(y)$ denote the set of points $x \in \mathbb{R}^2$ that are reachable from y. A nonempty set $S \subset \mathbb{R}^2$ is called *positive invariant* if $L(y) \subset S$ for every $y \in S$. This is equivalent to saying that

$$\Phi_i^t(x) \in S, \quad i \in \{0,1\},\ t \geq 0,\ x \in S.$$

Lemma 1. *The following statements hold.*

(1) The sets Γ and Γ° are positive invariant.
(2) We have $\Gamma^\circ \subset L(x)$ for every $x \in \mathbb{R}^2$.

PROOF: To simplify notation, we define

$$\gamma := \frac{\alpha}{\beta} > 1.$$

Fix a point $x \in \Gamma$. If $x = (0,0)$, the forward u_0 trajectory starting at x consists only of x and is therefore contained in Γ. If $x \neq (0,0)$ and if y is any point on the forward u_0 trajectory starting at x, we have

$$y_2^\gamma \leq x_1 x_2^{-\gamma} y_2^\gamma \leq (1-(1-x_2)^\gamma)\, x_2^{-\gamma} y_2^\gamma = \left(\frac{y_2}{x_2}\right)^\gamma - \left(\frac{y_2}{x_2} - y_2\right)^\gamma \leq 1-(1-y_2)^\gamma.$$

To obtain the last inequality, we used that

$$\frac{d}{dz}\left(z^\gamma - (z-y_2)^\gamma\right) = \gamma\left(z^{\gamma-1} - (z-y_2)^{\gamma-1}\right) > 0, \quad z > y_2,$$

and that $y_2/x_2 < 1$. Since $x_1 x_2^{-\gamma} y_2^\gamma = y_1$, it follows that the forward u_0 trajectory starting at x is contained in Γ. As $(1-x_1, 1-x_2) \in \Gamma$, we also have

$$\Phi_0^t(1-x_1, 1-x_2) \in \Gamma, \quad t > 0. \tag{10}$$

Equations (8) and (10) imply that the forward u_1 trajectory starting at x is contained in Γ as well. The proof for Γ° is analogous.

To prove the second statement, fix $x \in \mathbb{R}^2$ and $y \in \Gamma^\circ$. Since the critical points $(0,0)$ and $(1,1)$ are globally asymptotically stable, the set $L(x)$ contains a point z such that $z_2 \in (0,1)$. If $z_1 \geq z_2$, set

$$a(t) = \Phi_0^t(z), \quad t > 0.$$

We have

$$\left(e^{-\beta t} z_2\right)^\gamma < e^{-\alpha t} z_1. \tag{11}$$

Set

$$b(t) = e^{-\alpha t} z_1 + \left(1 - e^{-\beta t} z_2\right)^\gamma.$$

Since $\lim_{t\to\infty} b(t) = 1$ and since

$$b'(t) = -\alpha z_1 e^{-\alpha t} + \alpha z_2 e^{-\beta t}\left(1 - e^{-\beta t} z_2\right)^{\gamma-1} > 0$$

for t sufficiently large, we have $b(t) < 1$ for large t. This and (11) imply that $a(t) \in \Gamma^\circ$ for large t. In particular, there is $a \in L(x) \cap \Gamma^\circ$. If $z_1 < z_2$, set

$$a(t) = \Phi_0^t(1-z_1, 1-z_2), \quad t > 0.$$

As $1 - z_1 > 1 - z_2$, we have as before $a(t) \in \Gamma^\circ$ for large t. With (8) and symmetry of Γ° around $(\frac{1}{2}, \frac{1}{2})$, this yields

$$\Phi_1^t(z) \in \Gamma^\circ$$

for large t, so as in the case $z_1 \geq z_2$ there is $a \in L(x) \cap \Gamma^\circ$.

As $L(a) \subset L(x)$, we may assume without loss of generality that $x \in \Gamma^\circ$. Now, it suffices to show that one of the following statements holds, as this means there are $s, t > 0$ and $i \in \{0,1\}$ such that $y = \Phi_i^{(s,t)}(x)$.

(a) There is $\eta \in (0, \min\{x_2, y_2\}]$ such that

$$g(\eta) := 1 - (1-y_1)(1-y_2)^{-\gamma}(1-\eta)^\gamma - x_1 x_2^{-\gamma}\eta^\gamma = 0.$$

(b) There is $\eta \in [\max\{x_2, y_2\}, 1)$ such that

$$h(\eta) := 1 - (1-x_1)(1-x_2)^{-\gamma}(1-\eta)^\gamma - y_1 y_2^{-\gamma}\eta^\gamma = 0.$$

It is easy to see that $g(0)$, $g(1)$, $h(0)$, and $h(1)$ are negative. Since g and h are continuous, it is then enough to show that $g(\min\{x_2, y_2\}) \geq 0$ or $h(\max\{x_2, y_2\}) \geq 0$. First, assume that $x_2 \leq y_2$. If

$$(1 - y_1)(1 - y_2)^{-\gamma} \leq (1 - x_1)(1 - x_2)^{-\gamma},$$

we have

$$g(x_2) = 1 - (1 - y_1)(1 - y_2)^{-\gamma}(1 - x_2)^{\gamma} - x_1 \geq 0.$$

And if

$$(1 - x_1)(1 - x_2)^{-\gamma} \leq (1 - y_1)(1 - y_2)^{-\gamma},$$

we have

$$h(y_2) = 1 - (1 - x_1)(1 - x_2)^{-\gamma}(1 - y_2)^{\gamma} - y_1 \geq 0.$$

It remains to consider the case $x_2 > y_2$. If $x_1 x_2^{-\gamma} \leq y_1 y_2^{-\gamma}$, we have $g(y_2) \geq 0$. And if $y_1 y_2^{-\gamma} \leq x_1 x_2^{-\gamma}$, we have $h(x_2) \geq 0$. □

Since the semigroup $(\mathsf{P}^t)_{t\geq 0}$ is Feller and since Γ is compact and positive invariant, $(\mathsf{P}^t)_{t\geq 0}$ admits an invariant probability measure. In fact, as we will show below, the invariant probability measure is unique and absolutely continuous. For $x \in \mathbb{R}^2$, let

$$U(x) = (u_1(x), u_0(x))$$

be the (2×2) matrix whose first column is $u_1(x)$ and whose second column is $u_0(x)$. As stated in the following easily verified lemma, u_0 and u_1 are transversal at every point except for points on the line $x_1 = x_2$.

Lemma 2. *Let $x \in \mathbb{R}^2$. Then,* $\det U(x) = \alpha\beta(x_1 - x_2)$. *In particular,* $\det U(x) = 0$ *if and only if $x_1 = x_2$.*

In light of Lemmas 1 and 2, there are points (namely every point in Γ° not located on $x_1 = x_2$) that are reachable from every starting point in $\mathbb{R}^2$ and where u_0 and u_1 are transversal. By Theorem 1 in [1] or by Theorem 4.5 in [5], the invariant probability measure of $(\mathsf{P}^t)_{t\geq 0}$ is unique and absolutely continuous with respect to the product of Lebesgue measure on $\mathbb{R}^2$ and counting measure on $\{0, 1\}$.

Alternatively, one can follow the reasoning in [15], which leverages the contractive nature of the system. This is particularly simple in this case as the flows are deterministically uniformly contracting. Fixing two initial condition $x, y \in \mathbb{R}^2$ and $i \in \{0, 1\}$, we set $I_0 = i$ and let $X_t(x)$ and $X_t(y)$ be the solution to (9) with the same I_t process (and hence the same jump times) but starting initially from x and y respectively. If we define $r_t := X_t(x) - X_t(y)$, observe that

$$\|r_t\| \leq \|x - y\| \exp(-(\alpha \wedge \beta)t), \tag{12}$$

where $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^2$. Now let $f\colon \mathbb{R}^2 \times \{0, 1\} \to \mathbb{R}$ be an arbitrary test function which is 1-Lipschitz continuous, i.e. $|f(x, i) - f(y, j)| \leq \|x - y\| + \mathbb{1}_{i\neq j}$ for every $x, y \in \mathbb{R}^2$ and $i, j \in \{0, 1\}$. Since $\sup_{t\geq 0} \|X_t(x)\| < \infty$ for every $x \in \mathbb{R}^2$, the supremum of $(\mathsf{P}^t f)(x, i) - (\mathsf{P}^t f)(y, i) = \mathbf{E} f(X_t(x), I_t) - \mathbf{E} f(X_t(y), I_t)$ over all such test functions is equal to the 1-Wasserstein distance between $\mathsf{P}^t(x, i;\ \cdot\ )$ and $\mathsf{P}^t(y, i;\ \cdot\ )$. Denoting this distance by $\|\mathsf{P}^t(x, i;\ \cdot\ ) - \mathsf{P}^t(y, i;\ \cdot\ )\|_{W_1}$ and recalling that $\|\delta_{x,i} - \delta_{y,i}\|_{W_1} = \|x - y\|$ produces

$$\|\mathsf{P}^t(x, i;\ \cdot\ ) - \mathsf{P}^t(y, i;\ \cdot\ )\|_{W_1} \leq \|\delta_{x,i} - \delta_{y,i}\|_{W_1} e^{-(\alpha\wedge\beta)t}.$$

A simple coupling argument using the definition of the 1-Wasserstein distance as the infimum over all couplings and the resulting convexity of the 1-Wasserstein distance produces

$$\|\mu \mathsf{P}^t - \nu \mathsf{P}^t\|_{W_1} \le \|\mu - \nu\|_{W_1} e^{-(\alpha\wedge\beta)t} \tag{13}$$

for arbitrary initial probability measures μ and ν of bounded support. Observe in addition that there is a bounded positive invariant subset B of $\mathbb{R}^2$ (e.g., the set $[-1,2]^2$) with the following property: For every $R > 0$ there is $C > 0$ such that every switching trajectory starting from a point of distance less than R from the origin enters the set B in a time less than C. As a result, every invariant probability measure has bounded support, so the estimate in (13) proves in particular uniqueness of the invariant probability measure. Additionally, this proves a spectral gap for the Markov kernel P^t in the 1-Wasserstein distance which is independent of the switching rates. The estimate in (12) can also be used to show that the system has a random attractor which consists of a single point and that all of the Lyapunov exponents are negative. See [15] for more discussions in this direction.

To finish the proof of Proposition 1, it remains to show the statement about the support of the marginals μ_0 and μ_1. Since Γ is positive invariant and compact, the support of μ_0 and μ_1 is contained in Γ. And since Γ° is an open subset of the set of points $y \in \Gamma$ that are reachable from every starting point in Γ, Γ° is a subset of the support of μ_0 and μ_1, see e.g. [1, Lemma 6]. As the support of μ_0 and μ_1 is a closed set, it is necessarily equal to Γ.

5. Integral equations for invariant densities and CDF's

Recall that for $i \in \{0,1\}$, Ψ_i denotes the inverse flow associated with the vector field u_i. Lemma 2 in [3] implies that

$$\rho_i(x) = \int_{\mathbb{R}_+} \lambda_{1-i} e^{-\lambda_i t} \det \nabla_x \Psi_i^t(x) \rho_{1-i}(\Psi_i^t(x))\, dt, \quad i \in \{0,1\}. \tag{14}$$

Written in terms of the cumulative distribution functions (CDF's)

$$G_i(x_1,x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} \rho_i(y_1,y_2)\, dy_1\, dy_2, \quad (x_1,x_2)\in\mathbb{R}^2, \tag{15}$$

the integral equations in (14) become

$$G_i(x) = \int_{\mathbb{R}_+} \lambda_{1-i} e^{-\lambda_i t} G_{1-i}(\Psi_i^t(x))\, dt, \quad i \in \{0,1\}. \tag{16}$$

This is because for $i \in \{0,1\}$ and for any fixed $x \in \mathbb{R}^2$, $t \mapsto \Phi_i^t(x)$ is monotone in both components. Note that the integral on the right side of (16) can be rewritten as

$$\frac{\lambda_{1-i}}{\lambda_i} \mathbf{E} G_{1-i}(\Psi_i^T(x)),$$

where T is an exponential random variable with intensity λ_i.

Next, we generalize the integral equations in (14) by considering the evolution of (X,I) leading to the current state not just since the latest switch but over the latest n switches, $n \in \mathbb{N}$. For $i \in \{0,1\}$, $n \in \mathbb{N}$, $\mathbf{t} \in \mathbb{R}_+^n$, and $x \in \Gamma^\circ$, we define the Jacobian

$$J_i^{\mathbf{t}}(x) = \det \nabla_x \Psi_i^{\mathbf{t}}(x).$$

For $i \in \{0,1\}$, $n \in \mathbb{N}$, and $x \in \Gamma^\circ$, let

$$T_i^n(x) = \left\{ \mathbf{t} \in \mathbb{R}_+^n : \Psi_i^{\mathbf{t}}(x) \in \Gamma^\circ \right\}. \tag{17}$$

For $n \in \mathbb{N}$ and real-valued integrable functions h on Γ°, we define the transfer operator

$$\mathcal{Q}_n h(x) = \int_{T_0^n(x)} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t} \rangle} J_0^{\mathbf{t}}(x) h(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t}, \quad x \in \Gamma^\circ, \tag{18}$$

where

$$\lambda_i(n) = \begin{cases} \lambda_{1-i}\lambda_i \ldots \lambda_i, & n \equiv 0 \mod 2, \\ \lambda_{1-i}\lambda_i \ldots \lambda_{1-i}, & n \equiv 1 \mod 2 \end{cases} \tag{19}$$

is an alternating product of λ_{1-i} and λ_i with exactly n factors, and where

$$\lambda_i^{(n)} = \begin{cases} (\lambda_{1-i}, \lambda_i, \ldots, \lambda_i)^\top, & n \equiv 0 \mod 2, \\ (\lambda_i, \lambda_{1-i}, \ldots, \lambda_i)^\top, & n \equiv 1 \mod 2 \end{cases} \tag{20}$$

is a vector of length n whose components alternate between λ_i and λ_{1-i}. Then,

$$\rho_0 = \begin{cases} \mathcal{Q}_n \rho_0, & n \equiv 0 \mod 2, \\ \mathcal{Q}_n \rho_1, & n \equiv 1 \mod 2, \end{cases} \tag{21}$$

which can be deduced by iteratively plugging instances of (14) into one another and using the fact that the pushforward of a function under the cumulative flow $\Phi_i^{\mathbf{t}}$ is the composition of pushforwards under the individual flows $\Phi_i^{t_n}, \Phi_{1-i}^{t_{n-1}}, \ldots$

Remark 2. The formula in (21) can be generalized to switching systems with state space $U \times S$, where U is an open subset of $\mathbb{R}^n$ and S is a finite index set corresponding to a collection of smooth vector fields u on $\mathbb{R}^n$ that leave U positive invariant and are integrable, i.e. for any $x_0 \in U$ the initial-value problem $\dot{x} = u(x), x(0) = x_0$ has a unique solution, and this solution is defined for all $t \in \mathbb{R}$. Suppose the corresponding Markov semigroup admits an absolutely continuous invariant measure μ with invariant densities $(\rho_i)_{i \in S}$. For $i, j \in S$, let λ_i be the rate of switching away from vector field u_i and let $\lambda_{j,i}$ be the rate of switching from u_j to u_i. For $n \in \mathbb{N}$, $\mathbf{i} = (i_1, \ldots, i_n) \in S^n$ and for real-valued integrable functions h on U, define

$$\mathcal{Q}_{\mathbf{i}} h(x) = \int_{T_{\mathbf{i}}(x)} \prod_{j=2}^{n} \lambda_{i_{j-1}, i_j} e^{-\sum_{j=1}^n \lambda_{i_j} t_j} \Phi_{\mathbf{i}}^{\mathbf{t}} \# h(x)\, d\mathbf{t}, \quad x \in U,$$

where $\Phi_{\mathbf{i}}^{\mathbf{t}} = \Phi_{i_n}^{t_n} \circ \ldots \circ \Phi_{i_1}^{t_1}$, $T_{\mathbf{i}}(x) = \{\mathbf{t} \in \mathbb{R}_+^n : (\Phi_{\mathbf{i}}^{\mathbf{t}})^{-1}(x) \in U\}$ and where $\Phi_{\mathbf{i}}^{\mathbf{t}} \# h$ denotes the pushforward of h under $\Phi_{\mathbf{i}}^{\mathbf{t}}$. Then, we have for $n \in \mathbb{N}$ and $i_n \in S$ that

$$\rho_{i_n} = \sum_{i_{n-1} \neq i_n} \ldots \sum_{i_0 \neq i_1} \lambda_{i_0, i_1} \mathcal{Q}_{(i_1, \ldots, i_n)} \rho_{i_0}.$$

6. Singularities of the invariant densities

In this section, we prove Theorem 1 on singularities of the invariant density ρ_0 for slow switching. We first show that if $\lambda_0 < \alpha + \beta$, then ρ_0 is unbounded in any neighborhood of $(0, 0)$.

PROOF OF THEOREM 1, PART (1): Recall that the cumulative distribution function G_0 was defined in (15). Assuming that the invariant density ρ_0 is bounded by a constant C in some neighborhood of $(0,0)$, we conclude that

$$G_0(\epsilon^\alpha, \epsilon^\beta) < C\epsilon^{\alpha+\beta}$$

for sufficiently small ϵ. On the other hand, (16) implies that, for every $\epsilon > 0$,

$$\begin{aligned} G_0(\epsilon^\alpha, \epsilon^\beta) &= \int_0^\infty \lambda_1 e^{-\lambda_0 t} G_1(e^{\alpha t}\epsilon^\alpha, e^{\beta t}\epsilon^\beta)\ dt \\ &\geq \int_{\ln(1/\epsilon)}^\infty \lambda_1 e^{-\lambda_0 t} G_1(e^{\alpha t}\epsilon^\alpha, e^{\beta t}\epsilon^\beta)\ dt. \end{aligned} \tag{22}$$

For $t \geq \ln(\frac{1}{\epsilon})$, we have $e^{\beta t}\epsilon^\beta \geq 1$ and thus, as the support Γ of μ_1 is contained in $[0,1]^2$, we have $G_1(e^{\alpha t}\epsilon^\alpha, e^{\beta t}\epsilon^\beta) = G_1(+\infty, +\infty)$. Hence, the integral in the second line of (22) equals

$$G_1(+\infty,+\infty)\int_{\ln(1/\epsilon)}^\infty \lambda_1 e^{-\lambda_0 t}\ dt = G_1(+\infty,+\infty)\frac{\lambda_1}{\lambda_0}\epsilon^{\lambda_0}.$$

Since ϵ is arbitrary, we conclude that $\lambda_0 \geq \alpha + \beta$. □

The idea behind the above proof is very general and can be used with minor modifications to study existence and character of singularities in various other situations including high-dimensional ones. However, there is another interesting proof specific to the concrete vector fields u_0, u_1 we consider.

ANOTHER PROOF OF THEOREM 1, PART (1): Let $\epsilon \in (0,1)$. Since the point $(\epsilon^\alpha, \epsilon^\beta) \in \mathbb{R}^2$ is on $\partial\Gamma_0$, the left boundary curve of Γ, and since ρ_0 is identically zero outside of Γ, we have

$$G_0(\epsilon^\alpha, \epsilon^\beta) = \int_{-\infty}^{\epsilon^\alpha}\int_{-\infty}^{\infty} \rho_0(y_1,y_2)\, dy_1\, dy_2 = G_0(\epsilon^\alpha, +\infty).$$

Since our system can be viewed as a product of non-iteracting components, the marginal distribution

$$E \mapsto \int_E \int_{-\infty}^\infty \rho_0(y_1,y_2)\ dy_1\ dy_2$$

coincides with the stationary distribution of the one-dimensional system given by switching at rates λ_0 and λ_1 between the one-dimensional vector fields

$$v_0(x) = -\alpha x, \quad v_1(x) = -\alpha(x-1).$$

By Proposition 3.12 in [9], this distribution is a beta distribution with parameters $(\lambda_0/\alpha, \lambda_1/\alpha{+}1)$, so assuming that the invariant density ρ_0 is bounded by a constant C in a neigborhood of $(0,0)$, we obtain for small ϵ

$$cB(\epsilon^\alpha; \lambda_0/\alpha, \lambda_1/\alpha+1) = G_0(\epsilon^\alpha, \epsilon^\beta) \leq C\epsilon^{\alpha+\beta}, \tag{23}$$

where c is a normalizing constant and where $B(x;a,b)$ is the incomplete beta function

$$B(x;a,b) = \int_0^x t^{a-1}\,(1-t)^{b-1}\,dt.$$

Now if $\lambda_0 < \alpha + \beta$, it is straightforward to check that

$$\epsilon^{-(\alpha+\beta)}B(\epsilon^\alpha; \lambda_0/\alpha, \lambda_1/\alpha+1) \to \infty, \quad \text{as } \epsilon \to 0.$$

Thus, dividing (23) by $\epsilon^{\alpha+\beta}$ and taking $\epsilon \to 0$ completes the proof. $\square$

Now we prove the second part of Theorem 1, which asserts that ρ_0 blows up along the entire left boundary curve of Γ in case $\lambda_1 < \beta$.

Proof of Theorem 1, part (2): Let us introduce the functions

$$\phi^{(1)}(t) = e^{-\alpha t}, \quad \phi^{(2)}(t) = e^{-\beta t}, \quad \phi^{(2)}_\epsilon(t) = (1-\epsilon^\beta)e^{-\beta t}, \quad t \in \mathbb{R},\ \epsilon > 0,$$

so that $\{(\phi^{(1)}(t), \phi^{(2)}(t))\}_{t>0}$ is the forward u_0 trajectory starting at $(1,1)$ — the left boundary curve of Γ. Also, $\Phi_0^t(1, 1-\epsilon^\beta) = (\phi^{(1)}(t), \phi^{(2)}_\epsilon(t))$.

For $z > 0$, we define $t(z) = (\phi^{(1)})^{-1}(z) = -\ln(z)/\alpha$ and the open interval $I_\epsilon(z) = (\phi^{(2)}_\epsilon(t(z)), \phi^{(2)}(t(z))) \subset \mathbb{R}$. Also, for any set $I \subset \mathbb{R}$, we introduce

$$R_\epsilon(I) = \left\{(x_1, x_2) : x_1 \in \phi^{(1)}(I),\ x_2 \in I_\epsilon(x_1)\right\}.$$

Note that for all $t \in \mathbb{R}$ and all $t_1, t_2 \in \mathbb{R}$ satisfying $t_1 < t_2$, $\Phi_0^t : R_\epsilon(t_1, t_2) \to R_\epsilon(t_1 + t, t_2 + t)$ is a diffeomorphism. Also, for any $I, J \subset \mathbb{R}$,

$$R_\epsilon(I) \cap R_\epsilon(J) = R_\epsilon(I \cap J). \tag{24}$$

Suppose that there is $t > 0$ such that the invariant density ρ_0 is bounded by a constant C in a neighborhood of $\Phi_0^t(1,1)$. For ϵ sufficiently small, the set $R_\epsilon(t, t+\epsilon^\alpha)$ is contained in this neighborhood. Using the diffeomorphism property of Φ_0^t mentioned above, it is easy to check that there is a number $C' > 0$ such that for small ϵ the Lebesgue measure of $R_\epsilon(t, t+\epsilon^\alpha)$ is bounded by $C'\epsilon^{\alpha+\beta}$. Therefore, $\mu_0(R_\epsilon(t, t+\epsilon^\alpha)) \le CC'\epsilon^{\alpha+\beta}$ for ϵ small.

Now we derive a lower bound for $\mu_0(R_\epsilon(t, t+\epsilon^\alpha))$. For small $\epsilon > 0$, $\phi^{(1)}(\epsilon^\alpha) < 1 - \frac{\alpha}{2}\epsilon^\alpha$. Therefore, $R_\epsilon(0, \epsilon^\alpha) \supset ((1-\frac{\alpha}{2}\epsilon^\alpha, 1) \times (1-\epsilon^\beta, 1)) \cap \Gamma^\circ$. As μ_1 is supported on Γ, this yields

$$\mu_1(R_\epsilon(0, \epsilon^\alpha)) \ge \int_{1-\frac{\alpha}{2}\epsilon^\alpha}^\infty \int_{1-\epsilon^\beta}^\infty \rho_1(y_1, y_2)\, dy_2\, dy_1 \ge G_1(+\infty, +\infty)\frac{\lambda_0}{\lambda_1}\epsilon^{\lambda_1}, \tag{25}$$

where the second inequality follows from the proof of part (1), with the roles of λ_0 and λ_1 reversed. Using (14), (24) and the observation that $\Phi_0^{-s} : R_\epsilon(t, t+\epsilon^\alpha) \to R_\epsilon(t-s, t-s+\epsilon^\alpha)$ is a diffeomorphism, we have

$$\begin{aligned}
\mu_0(R_\epsilon(t, t+\epsilon^\alpha)) &= \int_0^\infty \lambda_1 e^{-\lambda_0 s}\mu_1(R_\epsilon(t-s, t-s+\epsilon^\alpha))\, ds \\
&\ge \int_0^\infty \lambda_1 e^{-\lambda_0 s}\mu_1(R_\epsilon(t-s, t-s+\epsilon^\alpha) \cap R_\epsilon(0, \epsilon^\alpha))\, ds \\
&= \int_0^\infty \lambda_1 e^{-\lambda_0 s}\mu_1(R_\epsilon((t-s, t-s+\epsilon^\alpha) \cap (0, \epsilon^\alpha)))\, ds \\
&\ge \int_{t-\epsilon^\alpha}^{t+\epsilon^\alpha} \lambda_1 e^{-\lambda_0 s}\mu_1(R_\epsilon((t-s, t-s+\epsilon^\alpha) \cap (0, \epsilon^\alpha))\, ds \\
&\ge c\int_{t-\epsilon^\alpha}^{t+\epsilon^\alpha} \mu_1(R_\epsilon((t-s, t-s+\epsilon^\alpha) \cap (0, \epsilon^\alpha)))\, ds =: cA(\epsilon).
\end{aligned}$$

Here, $c > 0$ is a constant that doesn't depend on ϵ. To complete the estimate of $\mu_0(R_\epsilon(t, t+\epsilon^\alpha))$ from below, we use the lower bound on $\mu_1(R_\epsilon(0, \epsilon^\alpha))$ derived

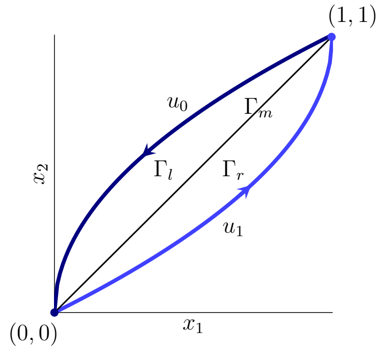

FIGURE 4. Illustration of the sets in (26).

in (25). We have

$$\begin{aligned}
A(\epsilon) &= \int_{t-\epsilon^\alpha}^{t+\epsilon^\alpha} \int_{\mathbb{R}^2} \mathbb{1}_{\{(x_1,x_2)\in R_\epsilon((t-s,t-s+\epsilon^\alpha)\cap(0,\epsilon^\alpha))\}}\ \mu_1(dx_1,dx_2)\ ds \\
&= \int_{t-\epsilon^\alpha}^{t+\epsilon^\alpha} \int_{\mathbb{R}^2} \mathbb{1}_{\{x_1\in\phi^{(1)}((t-s,t-s+\epsilon^\alpha)\cap(0,\epsilon^\alpha))\}} \mathbb{1}_{\{x_2\in I_\epsilon(x_1)\}}\ \mu_1(dx_1,dx_2)\ ds \\
&= \int_{\mathbb{R}^2} \mathbb{1}_{\{t(x_1)\in(0,\epsilon^\alpha)\}} \mathbb{1}_{\{x_2\in I_\epsilon(x_1)\}} \int_{t-\epsilon^\alpha}^{t+\epsilon^\alpha} \mathbb{1}_{\{t(x_1)\in(t-s,t-s+\epsilon^\alpha)\}}\ ds\ \mu_1(dx_1,dx_2) \\
&= \epsilon^\alpha \int_{\mathbb{R}^2} \mathbb{1}_{\{t(x_1)\in(0,\epsilon^\alpha)\}} \mathbb{1}_{\{x_2\in I_\epsilon(x_1)\}}\ \mu_1(dx_1,dx_2) \\
&= \epsilon^\alpha \mu_1(R_\epsilon(0,\epsilon^\alpha)) \geq G_1(+\infty,+\infty)\frac{\lambda_0}{\lambda_1}\epsilon^{\alpha+\lambda_1}.
\end{aligned}$$

In the next to last line we used that $t(x_1) \in (t-s, t-s+\epsilon^\alpha)$ if and only if $s \in (t-t(x_1), t-t(x_1)+\epsilon^\alpha)$, and if $t(x_1) \in (0,\epsilon^\alpha)$, one has $(t-t(x_1), t-t(x_1)+\epsilon^\alpha) \subset (t-\epsilon^\alpha, t+\epsilon^\alpha)$.

Combining the resulting lower bound for $\mu_0(R_\epsilon(t,t+\epsilon^\alpha))$ with the upper bound derived earlier, we see that for sufficiently small ϵ,

$$cG_1(+\infty,+\infty)\frac{\lambda_0}{\lambda_1}\epsilon^{\alpha+\lambda_1} \leq CC'\epsilon^{\alpha+\beta}.$$

This is possible only if $\lambda_1 \geq \beta$. □

7. Change of variables

The fixed-point equations in (21) give integral equations describing the invariant densities $\rho_i(x)$. These expressions were obtained by applying (14) multiple times. This captures the effect of pushing forward the density until the time of the nth switch. All of these expressions are written as integral operators where the integrals are taken over the length of the first n exponential times.

The goal in this section is to take the expressions for the pushforwards of the invariant densities in the case $n = 2$ as integrals over the switching times and change variables so that they can be viewed as integral operators over the state space Γ. The precise version is given in Lemma 4 at the end of this section.

It will be important to treat the points above, below and on the diagonal $\{x_1 = x_2\}$ differently. To this end, let us introduce the sets

$$\begin{aligned} \Gamma_l =&\{x \in \Gamma^\circ : x_1 < x_2\}, \qquad \Gamma_r = \{x \in \Gamma^\circ : x_1 > x_2\}, \\ \Gamma_m =&\{x \in \Gamma^\circ : x_1 = x_2\}, \qquad \Gamma_s = \Gamma_l \cup \Gamma_r, \end{aligned} \tag{26}$$

which are illustrated in Figure 7. Recalling the definition of T_i^n from (17), we define, for $i \in \{0,1\}$ and $x \in \Gamma^\circ$, the following sets of times when the first switch, going backwards in time, occurred respectively to the right and to the left of the diagonal:

$$\begin{aligned} R_i(x) &= \{(s,t) \in T_i^2(x) : \Psi_i^t(x) \in \Gamma_r\}, \\ L_i(x) &= \{(s,t) \in T_i^2(x) : \Psi_i^t(x) \in \Gamma_l\}\,. \end{aligned}$$

Observe that $R_1(x) = L_0((1,1)-x)$ and $L_1(x) = R_0((1,1)-x)$. The next lemma provides the regularity needed to perform the desired change of variables.

Lemma 3. *For any $x \in \Gamma^\circ$ and $i \in \{0,1\}$, $(s,t) \mapsto \Psi_i^{(s,t)}(x)$ is a diffeomorphism from $R_i(x)$ onto $\Psi_i^{R_i(x)}(x)$ and from $L_i(x)$ onto $\Psi_i^{L_i(x)}(x)$.*

Proof: It suffices to prove the statement for $i = 0$ because

$$\Psi_1^{(s,t)}(x) = (1,1) - \Psi_0^{(s,t)}((1,1)-x), \quad (s,t) \in T_1^2(x).$$

We only show that $(s,t) \mapsto \Psi_0^{(s,t)}(x)$ is a diffeomorphism on $R_0(x)$, as the proof for $L_0(x)$ is almost identical. We begin by showing that $(s,t) \mapsto \Psi_0^{(s,t)}(x)$ is injective on $R_0(x)$. To obtain a contradiction, suppose that this is not the case. Then, there exist two distinct vectors $(s_1,t_1),(s_2,t_2) \in R_0(x)$ such that $\Psi_0^{(s_1,t_1)}(x) = \Psi_0^{(s_2,t_2)}(x) =: y$. This implies that the u_0 trajectory through x and the u_1 trajectory through y intersect in two distinct points $z^{(1)} := \Psi_0^{t_1}(x)$ and $z^{(2)} := \Psi_0^{t_2}(x)$ that both lie in Γ_r. For any two points $x, y \in \Gamma^\circ$, the u_0 trajectory through x and the u_1 trajectory through y intersect in at most two distinct points, so $z^{(1)}$ and $z^{(2)}$ are the only points of intersection. Since $z^{(1)}, z^{(2)} \in \Gamma_r$, Lemma 2 implies that $\det U(z^{(i)}) > 0$ for $i \in \{0,1\}$. As a result, one trajectory crosses the other in the same direction at both points of intersection, which is impossible.

It remains to show that $\det \nabla_{(s,t)} \Psi_0^{(s,t)}(x) \neq 0$ for $(s,t) \in R_0(x)$. From (7), we derive

$$\Psi_0^{(s,t)}(x) = \begin{pmatrix} 1 - e^{\alpha s} + e^{\alpha(s+t)} x_1 \\ 1 - e^{\beta s} + e^{\beta(s+t)} x_2 \end{pmatrix}, \tag{27}$$

which yields

$$\nabla_{(s,t)} \Psi_0^{(s,t)}(x) = \begin{pmatrix} -\alpha e^{\alpha s} + \alpha x_1 e^{\alpha(s+t)} & \alpha x_1 e^{\alpha(s+t)} \\ -\beta e^{\beta s} + \beta x_2 e^{\beta(s+t)} & \beta x_2 e^{\beta(s+t)} \end{pmatrix}$$

and

$$\det \nabla_{(s,t)} \Psi_0^{(s,t)}(x) = \alpha\beta e^{(\alpha+\beta)s}(x_1 e^{\alpha t} - x_2 e^{\beta t}) > 0. \tag{28}$$

For the last inequality, we used that $\Psi_0^t(x) \in \Gamma_r$. □

Let $i \in \{0,1\}$ and $x \in \Gamma^\circ$. We denote the inverse of $(s,t) \mapsto \Psi_i^{(s,t)}(x)$ as a map from $R_i(x)$ onto $\Psi_i^{R_i(x)}(x)$ by $\chi_i^{r,x}$, and the inverse of $(s,t) \mapsto \Psi_i^{(s,t)}(x)$ as a map from $L_i(x)$ onto $\Psi_i^{L_i(x)}(x)$ by $\chi_i^{l,x}$. With $\lambda_i(2)$ and $\lambda_i^{(2)}$ defined as in (19) and (20), respectively, we also introduce the functions

$$\begin{aligned} f_i(\mathbf{t},x) =& \lambda_i(2)e^{-\langle\lambda_i^{(2)},\mathbf{t}\rangle}J_i^{\mathbf{t}}(x), \quad \mathbf{t}\in R_i(x)\cup L_i(x),\\ K_i^r(x,y) =& f_i(\chi_i^{r,x}(y),x)|\det\nabla_y\chi_i^{r,x}(y)|, \quad y\in\Psi_i^{R_i(x)}(x),\\ K_i^l(x,y) =& f_i(\chi_i^{l,x}(y),x)|\det\nabla_y\chi_i^{l,x}(y)|, \quad y\in\Psi_i^{L_i(x)}(x). \end{aligned}$$

We are now in a position to perform the desired change of variables on the operator $\mathcal{Q}_2$ which was defined in (18) and used in the fix point equations (21).

Lemma 4. *For any $i\in\{0,1\}$ and $x\in\Gamma^\circ$,*

$$\int_{R_i(x)}\lambda_i(2)e^{-\langle\lambda_i^{(2)},\mathbf{t}\rangle}J_i^{\mathbf{t}}(x)\rho_i(\Psi_i^{\mathbf{t}}x)\ d\mathbf{t} = \int_{\Psi_i^{R_i(x)}(x)}\rho_i(y)K_i^r(x,y)\ dy. \tag{29}$$

In (29), one can replace $R_i(x)$ and K_i^r together with $L_i(x)$ and K_i^l.

PROOF: The formula follows after applying the change of variables $y=\Psi_i^{\mathbf{t}}(x)$ justified by Lemma 3. □

8. Boundedness

In this section, we prove Theorem 2 that describes under which conditions the invariant density ρ_0 stays bounded. The main difficulties in proving this result stem from two sources: the exponential contraction in the vicinity of the critical points $(0,0)$ and $(1,1)$, and the fact, exhibited in Lemma 2, that the vector fields u_0 and u_1 are collinear at every point on the line $x_1=x_2$. As we saw in Section 6, exponential contraction is an essential problem that gives rise to singularities of the invariant densities for slow switching. The lack of ellipticity along the diagonal $x_1=x_2$ creates technical challenges because switches close to the diagonal have a less pronounced regularizing effect on the invariant densities. At the same time, switches close to the diagonal do not actively spoil the densities as long as they occur sufficiently far from the two critical points.

Throughout this section, we will use the following basic facts about the switching system, at times without explicitly referring to them.

Lemma 5. *For any $x\in\Gamma^\circ$, the following statements hold.*

(1) For any $i\in\{0,1\}$ there is a unique $\theta_i(x)\in\mathbb{R}$ such that $\det U(\Psi_i^{\theta_i(x)}x)=0$. We have $\Psi_i^{\theta_i(x)}(x)\in\Gamma^\circ$.

(2) We have

$$\frac{d}{dt}\det U(\Psi_i^t x)=\alpha\beta\left((i-x_2)\beta e^{\beta t}-(i-x_1)\alpha e^{\alpha t}\right),\quad i\in\{0,1\},$$

and

$$\frac{d}{dt}\det U\left(\Psi_0^t x\right)|_{t=0}>\beta\det U(x),\quad \frac{d}{dt}\det U\left(\Psi_1^t x\right)|_{t=0}<\alpha\det U(x).$$

In particular, if $x_1=x_2$ and if $\epsilon>0$, there is a unique $t_x(\epsilon)>0$ such that $\det U(\Psi_0^{t_x(\epsilon)}x)=\epsilon$.

(3) Suppose now that $x_1 = x_2$, and let $y \in \Gamma^\circ$ such that $x_1 < y_1 = y_2$. Then,

$$[\Psi_0^{t_y(\epsilon)}(y)]_2 - [\Psi_0^{t_x(\epsilon)}(x)]_2 > 0, \quad \epsilon > 0.$$

Here, $[z]_2 := z_2$ for $z = (z_1, z_2) \in \mathbb{R}^2$.

We omit the proof of this lemma. To illustrate our main strategy for establishing Theorem 2, we first show that ρ_0 is bounded on the part of Γ that lies below the diagonal $x_1 = x_2$. This statement has a comparatively simple proof as we do not need to address the lack of ellipticity and as there is no danger of entering regions with strong exponential contraction.

Proposition 2. *Let $K \subset \Gamma$ be a compact set such that $x_1 > x_2$ for all $x \in K$. Then, ρ_0 is bounded on K for every $\lambda_0, \lambda_1 > 0$.*

PROOF: Using (21) for $n = 2$ and the fact that $x \in \Gamma_r$ implies $T_0^2(x) = R_0(x)$, we have

$$\rho_0(x) = \int_{R_0(x)} \lambda_0(2) e^{-\langle \lambda_0^{(2)}, \mathbf{t} \rangle} J_0^{\mathbf{t}}(x) \rho_0(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t}, \quad x \in \Gamma_r. \tag{30}$$

Lemma 4 then yields

$$\rho_0(x) = \int_{\Psi_0^{R_0(x)}(x)} \rho_0(y) K_0^r(x, y)\, dy. \tag{31}$$

Since ρ_0 is integrable, it is enough to show that there is $c > 0$ such that

$$K_0^r(x, y) \le c, \quad x \in K \cap \Gamma_r,\ y \in \Psi_0^{R_0(x)}(x).$$

Let $x \in K \cap \Gamma_r$, $y \in \Psi_0^{R_0(x)}(x)$, and set $\mathbf{t} = (s, t) = \chi_0^{r,x}(y)$. Since $\Psi_0^t(x) \in \Gamma_r$, the matrix $U(\Psi_0^t x)$ is invertible. In [2], proof of Theorem 2, the formula

$$\nabla_x \Psi_0^{(s,t)}(x) = -\nabla_{(s,t)} \Psi_0^{(s,t)}(x) U(\Psi_0^t x)^{-1} \nabla_x \Psi_0^t(x)$$

was established. It shows how the effect of variations in the initial point on the final point after two switches (the lefthand side) can be translated into an equivalent variation in the switching times, with the translation given by the two rightmost terms on the righthand side. If we set $z = \Psi_0^t(x)$, the formula above yields

$$J_0^{(s,t)}(x) = \det \nabla_{(s,t)} \Psi_0^{(s,t)}(x) \det U(z)^{-1} \det \nabla_x \Psi_0^t(x).$$

Hence, since $\det \nabla_y \chi_0^{r,x}(y) = (\det \nabla_{(s,t)} \psi_0^{(s,t)}(x))^{-1}$,

$$K_0^r(x, y) = \lambda_0(2) e^{-\langle \lambda_0^{(2)}, \mathbf{t} \rangle} \det U(z)^{-1} \det \nabla_x \Psi_0^t(x) = \lambda_0(2) e^{-\lambda_1 s} \frac{e^{(\alpha+\beta-\lambda_0)t}}{\alpha\beta(z_1 - z_2)}.$$

By definition, t is the time it takes to move backward along the u_0 trajectory from x to z. As $z \in \Gamma^\circ$ and thus $z_2 < 1$, we have

$$t < -\frac{1}{\beta} \ln(x_2) \le \sup_{x \in K} \left(-\frac{1}{\beta} \ln(x_2) \right) < \infty,$$

where one should note that K is compact and only contains points $x = (x_1, x_2)$ such that $x_2 > 0$. As z lies on the backward u_0 trajectory starting at x, we also have

$$\frac{1}{z_1 - z_2} \le \frac{1}{x_1 - x_2} \le \sup_{x \in K} \frac{1}{x_1 - x_2} < \infty.$$

This completes the proof. □

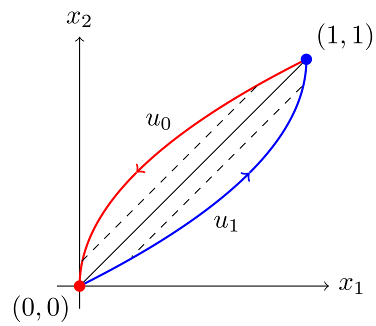

FIGURE 5. The dashed lines bound a strip around the diagonal, a region where switches have a less pronounced smoothing effect.

8.1. **Switches close to Γ_m.** If $y \in \Psi_0^{R_0(x)}(x)$ is chosen in such a way that z, the point where the switch from u_1 to u_0 occurs (cf. proof of Proposition 2), is close to the diagonal Γ_m, the term $(z_1 - z_2)^{-1}$ is very large, and so is $K_0^r(x, y)$. For $x \in \Gamma_l \cup \Gamma_m$, the point z can become arbitrarily close to Γ_m, which prevents $K_0^r(x, \cdot)$ from being bounded on $\Psi_0^{R_0(x)}(x)$. The proof of Proposition 2 can therefore not be extended to the case of $x \in \Gamma_l \cup \Gamma_m$ in a straightforward way. In this subsection, we describe an approach for dealing with this difficulty.

For any $n \in \mathbb{N}$, $x \in \Gamma^\circ$, and $\epsilon > 0$, let

$$M_n^\epsilon(x) = \left\{ \mathbf{t} \in T_0^n(x) : \left| \det U(\Psi_0^{(t_{n-j}, \ldots, t_n)} x) \right| < \epsilon, 0 \le j \le n-1 \right\},$$
$$S_n^\epsilon(x) = \left\{ \mathbf{t} \in T_0^{n+1}(x) : (t_3, \ldots, t_{n+1}) \in M_{n-1}^\epsilon(x), \left| \det U(\Psi_0^{(t_2, \ldots, t_{n+1})} x) \right| > \epsilon \right\}.$$

The condition $(t_3, \ldots, t_{n+1}) \in M_{n-1}^\epsilon(x)$ is void for $n = 1$. Observe that for $\mathbf{t} \in M_n^\epsilon(x)$, all points $\Psi_0^{(t_{n-j}, \ldots, t_n)}(x), 0 \le j \le n-1$, lie within a strip around the diagonal $x_1 = x_2$ whose width decreases linearly in ϵ (see Figure 5). For $\mathbf{t} \in S_n^\epsilon(x)$, the points $\Psi_0^{(t_{n+1-j}, \ldots, t_{n+1})}(x), 0 \le j \le n-2$, lie inside of the strip, while $\Psi_0^{(t_2, \ldots, t_{n+1})}(x)$ lies outside of it.

For $h \in L^1(\Gamma^\circ)$, we define

$$\mathcal{A}_n^\epsilon h(x) = \int_{S_n^\epsilon(x)} \lambda_0(n+1) e^{-\langle \lambda_0^{(n+1)}, \mathbf{t} \rangle} J_0^{\mathbf{t}}(x) h(\Psi_0^{\mathbf{t}} x) \, d\mathbf{t},$$
$$\mathcal{B}_n^\epsilon h(x) = \int_{M_n^\epsilon(x)} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t} \rangle} J_0^{\mathbf{t}}(x) h(\Psi_0^{\mathbf{t}} x) \, d\mathbf{t}.$$

To avoid distinguishing between the cases of even n and odd n, we introduce the shorthand

$$i_n = \begin{cases} 0, & n \equiv 0 \mod 2, \\ 1, & n \equiv 1 \mod 2. \end{cases}$$

Lemma 6. *For $n \in \mathbb{N}$, $x \in \Gamma^\circ$, and $\epsilon > 0$,*

$$\rho_0(x) = \sum_{k=1}^n \mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) + \mathcal{B}_n^\epsilon \rho_{i_n}(x).$$

PROOF: The proof is by induction. The formula in (21) gives

$$\rho_0(x) = \int_{T_0^1(x)} \lambda_1 e^{-\lambda_0 t} J_0^t(x) \rho_1(\Psi_0^t x)\ dt. \tag{32}$$

Statement (2) in Lemma 5 implies that for $y \in \mathbb{R}^2$, $t \in \mathbb{R}$, and $i \in \{0,1\}$,

$$\frac{d}{dt} \det U(\Psi_i^t y) = 0$$

if and only if

$$t = \frac{1}{\beta - \alpha} \ln \left(\frac{\alpha(y_1 - i)}{\beta(y_2 - i)} \right).$$

This shows that the set $T_0^1(x)$ is, up to a set of Lebesgue measure zero, the disjoint union of $M_1^\epsilon(x)$ and the set of $t \in \mathbb{R}_+$ such that $(s,t) \in S_1^\epsilon(x)$ for some $s \in \mathbb{R}_+$. Notice that the set $T_0^1(x)$ does not depend on ϵ, and that the decomposition $T_0^1(x) = M_1^\epsilon(x) \cup \{t \in \mathbb{R}_+ : \exists s \in \mathbb{R}_+ \text{ such that } (s,t) \in S_1^\epsilon(x)\}$, up to a set of Lebesgue measure zero, is valid for every choice of $\epsilon > 0$. The right side of (32) can thus be written as

$$\int_{t : \exists s \text{ s.t. } (s,t) \in S_1^\epsilon(x)} \lambda_1 e^{-\lambda_0 t} J_0^t(x) \rho_1(\Psi_0^t x)\ dt + \mathcal{B}_1^\epsilon \rho_1(x). \tag{33}$$

In complete analogy to (32), we have

$$\rho_1(y) = \int_{T_1^1(y)} \lambda_0 e^{-\lambda_1 s} J_1^s(y) \rho_0(\Psi_1^s y)\ ds. \tag{34}$$

If we plug this identity into the first summand in (33), we obtain the desired formula in the base case $n = 1$.

In the induction step, assume the formula holds for some $n \in \mathbb{N}$. With the notation $\mathbf{t} = (t_2, \dots, t_{n+1})$, we can write

$$\begin{aligned} &\mathcal{B}_n^\epsilon \rho_{i_n}(x) \\ &= \int_{M_n^\epsilon(x)} \int_{T_{i_n}^1(\Psi_0^{\mathbf{t}} x)} \lambda_0 (n+1) e^{-\langle \lambda_0^{(n+1)}, (t_1, \mathbf{t})^\top \rangle} J_0^{(t_1, \mathbf{t})}(x) \rho_{i_{n+1}}(\Psi_0^{(t_1,\mathbf{t})} x)\ dt_1\ d\mathbf{t}. \end{aligned} \tag{35}$$

The set

$$\left\{ (t_1, \mathbf{t}) \in \mathbb{R}_+^{n+1} : \mathbf{t} \in M_n^\epsilon(x),\ t_1 \in T_{i_n}^1(\Psi_0^{\mathbf{t}} x) \right\}$$

is, again up to a set of Lebesgue measure zero, the disjoint union of $M_{n+1}^\epsilon(x)$ and the set of $(t_1, \mathbf{t}) \in \mathbb{R}_+^{n+1}$ such that $(t_0, t_1, \mathbf{t}) \in S_{n+1}^\epsilon(x)$ for some $t_0 \in \mathbb{R}_+$. Therefore, the right side of (35) becomes

$$\mathcal{B}_{n+1}^\epsilon \rho_{i_{n+1}}(x) + \int_{\mathbf{t} : \exists t_0 \text{ s.t. } (t_0, \mathbf{t}) \in S_{n+1}^\epsilon(x)} \lambda_0 (n+1) e^{-\langle \lambda_0^{(n+1)}, \mathbf{t} \rangle} J_0^{\mathbf{t}}(x) \rho_{i_{n+1}}(\Psi_0^{\mathbf{t}} x)\ d\mathbf{t},$$

where we have set $\mathbf{t} = (t_1, \dots, t_{n+1})$. It remains to show that the integral term above equals $\mathcal{A}_{n+1}^\epsilon \rho_{i_{n+2}}$. This follows from plugging (32) or (34) into said integral term. □

Next, we show that for large n, the contribution of $\mathcal{B}_n^\epsilon \rho_{i_n}$ in the formula from Lemma 6 is small.

Lemma 7. *We have* $\lim_{n \to \infty} \|\mathcal{B}_n^\epsilon\|_{op} = 0$*, where* $\|\cdot\|_{op}$ *is the operator norm for operators on* $L^1(\Gamma^\circ)$*.*

PROOF: For fixed $n \in \mathbb{N}$ and $h \in L^1(\Gamma^\circ)$, we have

$$\|\mathcal{B}_n^\epsilon h\|_{L^1} \leq \int_{\widehat{M_n^\epsilon}} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} J_0^{\mathbf{t}}(x) |h(\Psi_0^{\mathbf{t}} x)| \, dx \, d\mathbf{t}, \tag{36}$$

where

$$\widehat{M_n^\epsilon} = \{(x, \mathbf{t}) \in \Gamma^\circ \times \mathbb{R}_+^n : \mathbf{t} \in M_n^\epsilon(x)\}.$$

Let us set

$$M_n^\epsilon = \bigcup_{x \in \Gamma^\circ} M_n^\epsilon(x).$$

Since $\widehat{M_n^\epsilon} \subset \Gamma^\circ \times M_n^\epsilon$, the right side of (36) is bounded by

$$\int_{M_n^\epsilon} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} \int_{\Gamma^\circ} J_0^{\mathbf{t}}(x) |h(\Psi_0^{\mathbf{t}} x)| \, dx \, d\mathbf{t}.$$

With the change of variables $y = \Psi_0^{\mathbf{t}}(x)$, the expression above becomes

$$\|h\|_{L^1} \int_{M_n^\epsilon} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} \, d\mathbf{t}.$$

Thus,

$$\|\mathcal{B}_n^\epsilon\|_{op} \leq \int_{M_n^\epsilon} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} \, d\mathbf{t} =: b_n.$$

As $M_{n+1}^\epsilon \subset \mathbb{R}_+ \times M_n^\epsilon$, we have

$$b_{n+1} \leq \int_{\mathbb{R}_+} \lambda_{i_{n+1}} e^{-\lambda_{i_n} t} \, dt \int_{M_n^\epsilon} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} \, d\mathbf{t} = \frac{\lambda_{i_{n+1}}}{\lambda_{i_n}} b_n,$$

so $\frac{b_{n+1}}{b_n}$ is bounded. To show that $\lim_{n\to\infty} b_n = 0$, it then suffices to show that there are $c \in (0,1)$ and $N \in \mathbb{N}$ such that

$$b_{2n+3} \leq c b_{2n+1}, \quad n \geq N.$$

We claim that there are $\tau_1, \tau_2, T > 0$ such that for n sufficiently large,

$$M_{2n+3}^\epsilon \subset ((\mathbb{R}_+ \times ((0, \tau_1] \cup [\tau_2, \infty))) \cup ((0, T) \times (\tau_1, \tau_2))) \times M_{2n+1}^\epsilon. \tag{37}$$

The idea behind this claim is the following: If the time between two consecutive switches that both happen close to the diagonal Γ_m is neither very short nor very long, i.e. if it falls within (τ_1, τ_2), then, at the time of the second switch, the switching trajectory cannot end up close to the critical point of the vector field to which the second switch is made. As a result, the time spent in Γ° after the second switch, following the time-reversed flow, is bounded by T. To prove the claim in (37), we first notice that for all $\tau > 0$

$$\det U(\Phi_1^\tau(0,0)) \neq 0.$$

Let us pick one such τ. Since the map

$$(x, t) \mapsto \det U(\Phi_1^t x)$$

is jointly continuous, there are $\hat{\epsilon} > 0$, $\tau_2 > \tau_1 > 0$ and a neighborhood V of $(0,0)$ such that

$$|\det U(\Phi_1^t x)| > \hat{\epsilon}, \quad t \in (\tau_1, \tau_2), \ x \in V.$$

Then, set

$$T = 1 + \sup\{t \geq 0 : \exists x \in V^c \cap \Gamma^\circ \text{ s.t. } \Psi_0^t(x) \in \Gamma^\circ\}.$$

This defines a finite quantity because $x_2 e^{\beta t} > 1$ for $t > -\frac{1}{\beta}\ln(x_2)$, and $\sup_{x\in V^c\cap\Gamma^\circ} -\ln(x_2) < \infty$. Let $(t_1,\ldots,t_{2n+3}) \in M^\epsilon_{2n+3}$. Then, $(t_1,\ldots,t_{2n+3}) \in M^\epsilon_{2n+3}(x)$ for some $x \in \Gamma^\circ$, and in particular $(t_3,\ldots,t_{2n+3}) \in M^\epsilon_{2n+1}(x)$. Suppose that $t_2 \in (\tau_1,\tau_2)$ and set

$$z = \Psi^{(t_2,\ldots,t_{2n+3})}(x).$$

We claim that $z \in V^c$. If z was an element of V, we would have

$$\hat\epsilon < |\det U(\Phi_1^{t_2} z)| = |\det U(\Psi_0^{(t_3,\ldots,t_{2n+3})}x)|.$$

On the other hand, as $(t_3,\ldots,t_{2n+3}) \in M^\epsilon_{2n+1}(x)$, we have

$$|\det U(\Psi_0^{(t_3,\ldots,t_{2n+3})}x)| < \epsilon \le \hat\epsilon$$

for ϵ sufficiently small, a contradiction.

From $z \in V^c$ it follows that $\Psi_0^t(z) \notin \Gamma^\circ$ for $t \ge T$. Hence, $t_1 < T$, and the inclusion in (37) is proved. As a result,

$$b_{2n+3} \le c b_{2n+1},$$

where

$$c = \int_{(0,\tau_1]\cup[\tau_2,\infty)} \lambda_1 e^{-\lambda_1 t}\ dt + \int_0^T \lambda_0 e^{-\lambda_0 s}\ ds \int_{\tau_1}^{\tau_2} \lambda_1 e^{-\lambda_1 t}\ dt < 1.$$

This completes the proof of the lemma. □

Corollary 1. *For $\epsilon > 0$ sufficiently small, we have $\lim_{n\to\infty} \mathcal{B}^\epsilon_n \rho_{i_n}(x) = 0$ for Lebesgue almost every $x \in \Gamma^\circ$.*

Proof: Recall from the proof of Lemma 6 that

$$\begin{aligned}&\mathcal{B}^\epsilon_n \rho_{i_n}(x)\\ =&\mathcal{B}^\epsilon_{n+1}\rho_{i_{n+1}}(x) + \int_{\mathbf{t}:\exists t_0 \text{ s.t. } (t_0,\mathbf{t})\in S^\epsilon_{n+1}(x)} \lambda_0(n+1) e^{-\langle \lambda_0^{(n+1)},\mathbf{t}\rangle} J_0^{\mathbf{t}}(x)\rho_{i_{n+1}}(\Psi_0^{\mathbf{t}} x)\ d\mathbf{t}.\end{aligned}$$

Thus, $(\mathcal{B}^\epsilon_n\rho_{i_n})_{n\ge1}$ is a pointwise monotone decreasing sequence of nonnegative functions. To prove the corollary, it is therefore enough to show that $(\mathcal{B}^\epsilon_n\rho_{i_n})_{n\ge1}$ converges to 0 in $L^1(\Gamma^\circ)$, which follows immediately from Lemma 7. □

The next lemma is a counterpart to the change of variables - formula in Lemma 4. For $i\in\{0,1\}$, $x\in\Gamma^\circ$, and $\epsilon > 0$, let

$$\begin{aligned}R_i^\epsilon(x) =&\{(s,t)\in T_i^2(x) : \det U(\Psi_i^t x) > \epsilon\},\\ L_i^\epsilon(x) =&\{(s,t)\in T_i^2(x) : \det U(\Psi_i^t x) < -\epsilon\}.\end{aligned}$$

In addition, we define

$$\begin{aligned}\mathcal{I}_i^{r,\epsilon}(x) &= \int_{\Psi_i^{R_i^\epsilon(x)}(x)} \rho_i(y) K_i^r(x,y)\ dy,\\ \mathcal{I}_i^{l,\epsilon}(x) &= \int_{\Psi_i^{L_i^\epsilon(x)}(x)} \rho_i(y) K_i^l(x,y)\ dy,\\ \mathcal{I}_i^{\epsilon}(x) &= \mathcal{I}_i^{r,\epsilon}(x) + \mathcal{I}_i^{l,\epsilon}(x).\end{aligned}$$

Lemma 8. *Let $x \in \Gamma^\circ$ and $\epsilon > 0$. Then,*

$$\mathcal{A}_1^\epsilon \rho_{i_2}(x) = \mathcal{I}_0^\epsilon(x).$$

For $k > 1$, we have

$$\mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) = \int_{M_{k-1}^\epsilon(x)} \lambda_0(k-1) e^{\langle (\alpha+\beta)\mathbb{1} - \lambda_0^{(k-1)}, \mathbf{t}\rangle} \mathcal{I}_{i_{k+1}}^\epsilon(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t}, \tag{38}$$

where $\mathbb{1} = (1, 1, \dots, 1)^\top$.

Proof: Observe that

$$\begin{aligned}\mathcal{A}_1^\epsilon \rho_{i_2}(x) = &\int_{R_0^\epsilon(x)} \lambda_0(2) e^{-\langle \lambda_0^{(2)}, \mathbf{t}\rangle} J_0^{\mathbf{t}}(x) \rho_{i_2}(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t} \\ &+ \int_{L_0^\epsilon(x)} \lambda_0(2) e^{-\langle \lambda_0^{(2)}, \mathbf{t}\rangle} J_0^{\mathbf{t}}(x) \rho_{i_2}(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t},\end{aligned} \tag{39}$$

and that Lemma 4 continues to hold if one replaces $R_0(x)$ with $R_0^\epsilon(x)$ and $L_0(x)$ with $L_0^\epsilon(x)$. The change of variables–formula in Lemma 4 then yields that the right side of (39) equals $\mathcal{I}_0^\epsilon(x)$.

For $k > 1$, write $\mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x)$ as

$$\begin{aligned}&\int_{M_{k-1}^\epsilon(x)} dt_3 \dots dt_{k+1}\ \lambda_0(k-1) e^{-\langle \lambda_0^{(k-1)}, (t_3,\dots,t_{k+1})^\top\rangle} \\ &\Big(\int_{R_{i_{k+1}}^\epsilon(\Psi_0^{(t_3,\dots,t_{k+1})} x)} dt_1\ dt_2\ \lambda_{i_{k+1}}(2) e^{-\langle \lambda_{i_{k+1}}^{(2)}, (t_1,t_2)^\top\rangle} J_0^{\mathbf{t}}(x) \rho_{i_{k+1}}(\Psi_0^{\mathbf{t}} x) \\ &+ \int_{L_{i_{k+1}}^\epsilon(\Psi_0^{(t_3,\dots,t_{k+1})} x)} dt_1\ dt_2\ \lambda_{i_{k+1}}(2) e^{-\langle \lambda_{i_{k+1}}^{(2)}, (t_1,t_2)^\top\rangle} J_0^{\mathbf{t}}(x) \rho_{i_{k+1}}(\Psi_0^{\mathbf{t}} x)\Big),\end{aligned} \tag{40}$$

where $\mathbf{t} = (t_1, \dots, t_{k+1})^\top$. Since

$$J_0^{\mathbf{t}}(x) = J_{i_{k+1}}^{(t_1,t_2)}(\Psi_0^{(t_3,\dots,t_{k+1})} x) J_0^{(t_3,\dots,t_{k+1})}(x) = J_{i_{k+1}}^{(t_1,t_2)}(\Psi_0^{(t_3,\dots,t_{k+1})} x) e^{(\alpha+\beta)\sum_{j=3}^{k+1} t_j},$$

we obtain the desired formula after applying Lemma 4 to the integrals in the second and third line of (40). □

Lemma 9. *For $i \in \{0, 1\}$, $\epsilon > 0$, and $x \in \Gamma^\circ$, we have*

$$\mathcal{I}_i^{r,\epsilon}(x) \le \begin{cases} \frac{\lambda_0\lambda_1}{\epsilon}, & \lambda_i \ge \alpha+\beta, \\ \frac{\lambda_0\lambda_1}{\epsilon} e^{(\alpha+\beta-\lambda_i)\tau_i(x)}, & \lambda_i < \alpha+\beta, \end{cases}$$

where

$$\tau_i(x) = \sup\{t \ge 0 : \Psi_i^t(x) \in \Gamma^\circ\}.$$

The estimate continues to hold if one replaces $\mathcal{I}_i^{r,\epsilon}(x)$ with $\mathcal{I}_i^{l,\epsilon}(x)$.

Proof: Setting $(s, t) = \chi_i^{r,x}(y)$ and $z = \Psi_i^t(x)$, we have

$$K_i^r(x, y) = \lambda_i(2) e^{-\langle \lambda_i^{(2)}, (s,t)^\top\rangle} |\det U(z)^{-1}| \det \nabla_x \Psi_i^t(x) = \lambda_i(2) e^{-\lambda_{1-i} s} \frac{e^{(\alpha+\beta-\lambda_i)t}}{|\det U(z)|}.$$

Since $|\det U(z)| \ge \epsilon$ for $y \in \Psi_i^{R_i^\epsilon(x)}(x)$, the desired estimate follows. □

For $n \in \mathbb{N}$, $\epsilon > 0$, $x \in \Gamma^\circ$, and $\lambda_0, \lambda_1 > 0$, set

$$\mathcal{M}_n^{\epsilon,\lambda_0,\lambda_1}(x) = \int_{M_n^\epsilon(x)} e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} \, d\mathbf{t}.$$

Lemma 10. *For any $\lambda_0, \lambda_1 > 0$, there is a function $f(\epsilon)$ such that $\lim_{\epsilon \downarrow 0} f(\epsilon) = 0$ and*

$$\mathcal{M}_{n+2}^{\epsilon,\lambda_0,\lambda_1}(x) \le f(\epsilon) \mathcal{M}_n^{\epsilon,\lambda_0,\lambda_1}(x), \quad n \in \mathbb{N}, \; x \in \Gamma^\circ.$$

Proof: For $\epsilon > 0$ and $i \in \{0,1\}$, let

$$\Gamma_i^\epsilon = \left\{ x \in \Gamma^\circ : |\det U(x)| < \epsilon, \; (-1)^i x_1 > \frac{1}{2} - i \right\}.$$

For $n \in \mathbb{N}$, $x \in \Gamma^\circ$, and $\epsilon > 0$, we have

$$\mathcal{M}_{n+2}^{\epsilon,\lambda_0,\lambda_1}(x) = \int_{M_n^\epsilon(x)} e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} \sum_{(i,j)\in\{0,1\}^2} I_{i,j}^{\mathbf{t}} \, d\mathbf{t},$$

where

$$I_{i,j}^{\mathbf{t}} = \int_{t_2 \in \mathbb{R}_+ : \; \Psi_0^{(t_2,\mathbf{t})}(x) \in \Gamma_j^\epsilon} dt_2 \; e^{-\lambda_{i_n} t_2} \int_{t_1 \in \mathbb{R}_+ : \; \Psi_0^{(t_1,t_2,\mathbf{t})}(x) \in \Gamma_i^\epsilon} dt_1 \; e^{-\lambda_{i_{n+1}} t_1}.$$

For $\mathbf{t} \in M_n^\epsilon(x)$ and $(i,j) \in \{0,1\}^2$, we now derive an upper bound on $I_{i,j}^{\mathbf{t}}$. For $x \in \Gamma_0^\epsilon$, $t \ge 0$, and $\epsilon < \frac{\alpha}{2}(\alpha - \beta)$, part (2) of Lemma 5 yields

$$\begin{aligned} \frac{d}{dt} \det U(\Psi_0^t x) <& e^{\beta t} \beta(-\alpha(\alpha-\beta)x_1 - \det U(x)) \\ \le& e^{\beta t} \beta \left(-\frac{\alpha}{2}(\alpha-\beta) + \epsilon\right) \le \beta \left(-\frac{\alpha}{2}(\alpha-\beta) + \epsilon\right). \end{aligned}$$

Similarly,

$$\frac{d}{dt} \det U(\Psi_1^t x) > \beta \left(\frac{\alpha}{2}(\alpha-\beta) - \epsilon\right), \quad x \in \Gamma_1^\epsilon, \; t \ge 0.$$

Thus, we have for $x \in \Gamma_i^\epsilon$ and

$$t \ge \frac{2\epsilon}{\beta(\frac{\alpha}{2}(\alpha-\beta) - \epsilon)} =: \tau_\epsilon$$

the estimate

$$\begin{aligned} |\det U(\Psi_i^t x)| \ge& |\det U(\Psi_i^t x) - \det U(x)| - |\det U(x)| \\ >& \beta \left(\frac{\alpha}{2}(\alpha-\beta) - \epsilon\right) t - \epsilon \ge \epsilon. \end{aligned} \tag{41}$$

We distinguish between two cases. Suppose first that $j = i_{n+1}$. With (41), we obtain for any $t_2 \in \mathbb{R}_+$ such that $\Psi_0^{(t_2,\mathbf{t})}(x) \in \Gamma_j^\epsilon$ the estimate

$$\int_{t_1 \in \mathbb{R}_+ : \Psi_0^{(t_1,t_2,\mathbf{t})}(x) \in \Gamma_i^\epsilon} e^{-\lambda_{i_{n+1}} t_1} \, dt_1 \le \int_0^{\tau_\epsilon} e^{-\lambda_{i_{n+1}} t_1} \, dt_1 \le \tau_\epsilon.$$

Then,

$$I_{i,j}^{\mathbf{t}} \le \tau_\epsilon \int_0^\infty e^{-\lambda_{i_n} t_2} \, dt_2 = \frac{\tau_\epsilon}{\lambda_{i_n}}.$$

If $j = i_n$, we first use the obvious estimate

$$I_{i,j}^{\mathbf{t}} \le \frac{1}{\lambda_{i_{n+1}}} \int_{t_2 \in \mathbb{R}_+ : \Psi_0^{(t_2,\mathbf{t})}(x) \in \Gamma_j^\epsilon} e^{-\lambda_{i_n} t_2} \, dt_2. \tag{42}$$

If $\{t_2 \in \mathbb{R}_+ : \Psi_0^{(t_2,\mathbf{t})}(x) \in \Gamma_j^\epsilon\} = \emptyset$, we have $I_{i,j}^{\mathbf{t}} = 0$. Otherwise,

$$\tau := \inf\{t_2 \in \mathbb{R}_+ : \Psi_0^{(t_2,\mathbf{t})}(x) \in \Gamma_j^\epsilon\}$$

is a number in $[0,\infty)$ such that

$$\Psi_0^{(t_2,\mathbf{t})}(x) \notin \Gamma_j^\epsilon, \quad t_2 < \tau.$$

The estimate in (41) then implies

$$\left|\det U\left(\Psi_0^{(\tau+t,\mathbf{t})}(x)\right)\right| \geq \epsilon, \quad t \geq \tau_\epsilon,$$

so for every $t \geq \tau_\epsilon$, we have $\Psi_0^{(\tau+t,\mathbf{t})}(x) \notin \Gamma_j^\epsilon$. As a result, the right side of (42) is less than

$$\frac{1}{\lambda_{i_{n+1}}} \int_\tau^{\tau+\tau_\epsilon} e^{-\lambda_{i_n} t_2}\, dt_2 \leq \frac{\tau_\epsilon}{\lambda_{i_{n+1}}}.$$

It follows that

$$\mathcal{M}_{n+2}^{\epsilon,\lambda_0,\lambda_1}(x) \leq \frac{4\tau_\epsilon}{\lambda_0 \wedge \lambda_1} \mathcal{M}_n^{\epsilon,\lambda_0,\lambda_1}(x).$$

□

8.2. **Proof of Theorem 2, part (1), for $\lambda_1 > \alpha+\beta$.** In this subsection, we prove part (1) of Theorem 2 in the case where $\lambda_1 > \alpha+\beta$. Under this additional assumption, we can give a simpler proof than in the general case treated in Subsection 8.3. In light of Lemma 6 and Corollary 1, it is enough to show that there is $c > 0$ such that for small $\epsilon > 0$,

$$\sum_{k=1}^{\infty} \mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) < c, \quad x \in \Gamma^\circ. \tag{43}$$

Let

$$m = \max\{(\lambda_0 - (\alpha+\beta))^{-1}, (\lambda_0 - (\alpha+\beta))^{-1}(\lambda_1 - (\alpha+\beta))^{-1}\}.$$

It is easy to see that

$$\max\left\{\mathcal{M}_1^{\epsilon,\lambda_0-(\alpha+\beta),\lambda_1-(\alpha+\beta)}(x), \mathcal{M}_2^{\epsilon,\lambda_0-(\alpha+\beta),\lambda_1-(\alpha+\beta)}(x)\right\} \leq m$$

for all $x \in \Gamma^\circ$. Lemma 10 implies

$$\mathcal{M}_k^{\epsilon,\lambda_0-(\alpha+\beta),\lambda_1-(\alpha+\beta)}(x) \leq f(\epsilon)^{\frac{k-2}{2}} m = f(\epsilon)^{-1} m f(\epsilon)^{\frac{k}{2}}, \quad k \in \mathbb{N}. \tag{44}$$

Combining (44) with Lemmas 8 and 9, we obtain

$$\sum_{k=1}^{\infty} \mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) \leq \frac{2\lambda_0\lambda_1}{\epsilon}\left(1 + f(\epsilon)^{-1} m \sum_{k=1}^{\infty} \lambda_0(k) f(\epsilon)^{\frac{k}{2}}\right), \quad x \in \Gamma^\circ.$$

If ϵ is so small that $(\lambda_0 \vee \lambda_1)\sqrt{f(\epsilon)} < 1$, the series on the right converges. This completes the proof.

8.3. **Proof of Theorem 2, part (1).** Let $\lambda_0 > \alpha+\beta$ and $\lambda_1 \in (\beta, \alpha+\beta]$ (recall that the case $\lambda_1 > \alpha+\beta$ has been taken care of in Subsection 8.2). For $n \in \mathbb{N}$, $\epsilon > 0$, $x \in \Gamma^\circ$, and $\lambda_0, \lambda_1 \in \mathbb{R}$, set

$$\widehat{\mathcal{M}}_n^{\epsilon,\lambda_0,\lambda_1}(x) = \int_{M_n^\epsilon(x)} e^{-\langle \lambda_0^{(n+1)}, (\tau_{i_n}(\Psi_0^{\mathbf{t}} x), \mathbf{t})\rangle} \, d\mathbf{t}, \tag{45}$$

where one should recall that $\tau_i(x)$ was defined in the statement of Lemma 9. We have

$$\sum_{k=1}^\infty \mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) \le \frac{2\lambda_0\lambda_1}{\epsilon}\left(1 + \sum_{k=1}^\infty \lambda_0(k) \widehat{\mathcal{M}}_k^{\epsilon,\lambda_0-(\alpha+\beta),\lambda_1-(\alpha+\beta)}(x)\right). \tag{46}$$

Now we need to estimate $\widehat{\mathcal{M}}_n^{\epsilon,\lambda_0,\lambda_1}$ for $\lambda_0 > 0$ and $\lambda_1 \in (-\alpha, 0]$.

Lemma 11. *There are functions $c(\epsilon)$ and $f(\epsilon)$ such that $\lim_{\epsilon\downarrow 0} f(\epsilon) = 0$ and*

$$\widehat{\mathcal{M}}_n^{\epsilon,\lambda_0,\lambda_1}(x) \le c(\epsilon) f(\epsilon)^n, \quad n \in \mathbb{N},\ x \in \Gamma^\circ,\ \lambda_0 > 0,\ \lambda_1 \in (-\alpha, 0]. \tag{47}$$

By Lemma 11, the right side of (46) is bounded by

$$\frac{2\lambda_0\lambda_1}{\epsilon}\left(1 + c(\epsilon)\sum_{k=1}^\infty (\lambda_0 f(\epsilon))^k\right),$$

which does not depend on x and is finite for ϵ so small that $\lambda_0 f(\epsilon) < 1$. To complete the proof of Theorem 2, part (1), it remains to show Lemma 11. We will do this at the end of this subsection.

For $i \in \{0,1\}$, $\epsilon > 0$ and $x \in \Gamma^\circ$ such that $|\det U(x)| < \epsilon$, let

$$\tau_i^\epsilon(x) = \sup\{t \ge 0 : |\det U(\Psi_i^t x)| < \epsilon\} \wedge \sup\{t \ge 0 : \Psi_i^t(x) \in \Gamma^\circ\}.$$

It is easy to see that $\tau_1^\epsilon(x) = \tau_0^\epsilon((1,1)-x)$. For integers $j \ge 2$, $i \in \{0,1\}$, and $\epsilon > 0$, we define

$$V_i^\epsilon(j) = \{x \in \Gamma^\circ : |\det U(x)| < \epsilon,\ \tau_i^\epsilon(x) \in (j-1, j]\}.$$

In order to deal with short exit times for the strip of points $x \in \Gamma^\circ$ such that $|\det U(x)| < \epsilon$, we also define

$$\begin{aligned} V_i^\epsilon(1) &= \{x \in \Gamma^\circ : |\det U(x)| < \epsilon,\ \tau_i^\epsilon(x) \in (\epsilon^{\frac{\beta}{\alpha}}, 1]\},\\ V_i^\epsilon(0) &= \{x \in \Gamma^\circ : |\det U(x)| < \epsilon,\ \tau_i^\epsilon(x) \in (0, \epsilon^{\frac{\beta}{\alpha}}]\}.\end{aligned}$$

In the following lemma, we analyze the interplay of exit times $\tau_0^\epsilon(x)$ and $\tau_1^\epsilon(x)$.

Lemma 12. *There exists a family of constants $(m^\epsilon(j))_{\epsilon>0, j\in\mathbb{N}}$ with the following properties.*

1. *For $\epsilon > 0$ sufficiently small and for any $j \in \mathbb{N}$, one has $\tau_i^\epsilon(x) < m^\epsilon(j)$ for every $i \in \{0,1\}$ and $x \in V_{1-i}^\epsilon(j)$.*
2. *One has*

$$\lim_{\epsilon\downarrow 0} \sum_{j=1}^\infty j e^{\nu j} m^\epsilon(j) = 0, \quad \nu \in [0, \alpha).$$

Proof: We first define the constants $(m^\epsilon(j))_{\epsilon>0, j\in\mathbb{N}}$ and then verify the asserted properties (1) and (2). In a first step, we show that there is $c > 0$ such that for every $x \in \Gamma^\circ$ with $x_1 < \frac{1}{2}$,

$$\tau_0^\epsilon((1,1)-x) < c x_1. \tag{48}$$

For $y \in \Gamma^\circ$, we have

$$(1-y_1)^{\frac{\beta}{\alpha}} > 1-y_2$$

and there is a unique $t^*(y) > 0$, given by

$$t^*(y) = \sup\{t \geq 0 : \Psi_0^t(y) \in \Gamma^\circ\},$$

such that

$$\left(1-e^{\alpha t^*(y)}y_1\right)^{\frac{\beta}{\alpha}} = 1-e^{\beta t^*(y)}y_2.$$

For fixed $y_1 \in (0,1)$, one observes that $t^*(y)$ is an increasing function of y_2: the larger y_2, the smaller is $1-y_2$, and the larger is the absolute value of the derivative

$$\frac{d}{dt}\left(1-e^{\beta t}y_2\right) = -\beta e^{\beta t}y_2,$$

i.e. the longer it takes the decreasing term $(1-e^{\alpha t}y_1)^{\frac{\beta}{\alpha}}$ to catch up with $1-e^{\beta t}y_2$. Now, consider the function

$$a(x_1,x_2,t) = 1-e^{\alpha t}(1-x_1) - \left|1-e^{\beta t}(1-x_2)\right|^{\frac{\alpha}{\beta}}, \quad (x_1,x_2,t) \in \mathbb{R}^3.$$

Since $\frac{\alpha}{\beta} > 1$, the function a is C^1 on $\mathbb{R}^3$. As $a(0,0,0) = 0$ and $\partial_t a(0,0,0) = -\alpha < 0$, the implicit function theorem implies that there is an open neighborhood U of $(0,0)$ and a C^1 function $b : U \to \mathbb{R}$ that is uniquely determined by $b(0,0) = 0$ and

$$a(x_1,x_2,b(x_1,x_2)) = 0, \quad (x_1,x_2) \in U.$$

The reason for defining a in the first place is that the function b induced by a satisfies

$$b(x) = t^*((1,1)-x), \quad x \in \Gamma^\circ \cap U.$$

Thus, for $x \in \Gamma^\circ$ sufficiently close to $(0,0)$, we have

$$\tau_0^\epsilon((1,1)-x) \leq t^*((1,1)-x) \leq b(x_1, 1-(1-x_1)^{\frac{\beta}{\alpha}}).$$

The rightmost expression in the chain of inequalities above is a C^1 function of x_1 for x_1 close to 0, and yields 0 when evaluated at $x_1 = 0$. The mean-value theorem then implies (48) for a suitable $c > 0$ and for $x \in \Gamma^\circ$ with x_1 sufficiently close to 0. As $\tau_0^\epsilon((1,1)-x)$ is bounded for $x \in \Gamma^\circ$ such that $x_1 < \frac{1}{2}$, we can extend (48) to the set of such x by choosing a larger c.

Let q be the unique real number greater than 1 such that $q - q^{\frac{\beta}{\alpha}} = 1$, and let $c' > 0$ be a constant such that

$$\frac{1}{e^{(\alpha-\beta)j}-1} < c' e^{(\beta-\alpha)j}, \quad j \geq 1.$$

Then, define

$$m^\epsilon(j) = \begin{cases} c\min\left\{qe^{-\alpha(j-1)}, \epsilon c' \frac{2}{\alpha\beta} e^{(\beta-\alpha)(j-1)}\right\}, & j \geq 2, \\ c\frac{2}{\alpha\beta(\alpha-\beta)}\epsilon^{1-\frac{\beta}{\alpha}}, & j = 1. \end{cases}$$

Fix $\nu \in [0,\alpha)$, and choose $\eta > 0$ such that $\alpha - \beta < \nu + \eta < \alpha$. Let $C > 0$ be so large that $j+1 \leq Ce^{\eta j}$ for all $j \in \mathbb{N}$. In addition, let $\gamma > 0$ be so small that

$\gamma(\nu+\eta+\beta-\alpha)<1$. As $e^{\nu+\eta-\alpha}<1$, we obtain for sufficiently small ϵ the estimate

$$\begin{aligned}
&\sum_{j=1}^{\infty} j e^{\nu j} m^{\epsilon}(j) \\
\leq & c e^{\nu}\left(\frac{2}{\alpha\beta(\alpha-\beta)}\epsilon^{1-\frac{\beta}{\alpha}} + C\epsilon c' \frac{2}{\alpha\beta}\sum_{j=1}^{\lfloor -\ln(\epsilon^{\gamma})\rfloor} e^{(\eta+\nu-\alpha+\beta)j} + Cq \sum_{j=\lfloor -\ln(\epsilon^{\gamma})\rfloor+1}^{\infty} e^{(\eta+\nu-\alpha)j}\right) \\
\leq & c e^{\nu}\left(\frac{2}{\alpha\beta(\alpha-\beta)}\epsilon^{1-\frac{\beta}{\alpha}} - \frac{2Cc'\gamma}{\alpha\beta}\ln(\epsilon)\epsilon^{1-\gamma(\nu+\eta+\beta-\alpha)} + Cq\frac{\epsilon^{\gamma(\alpha-\nu-\eta)}}{1-e^{\nu+\eta-\alpha}}\right),
\end{aligned}$$

and the right side converges to 0 as $\epsilon \downarrow 0$. This establishes property (2).

We now show property (1). For symmetry reasons, we can restrict ourselves to the case $i=1$, i.e. we will show that $\tau_1^{\epsilon}(x) < m^{\epsilon}(j)$ for all $x \in V_0^{\epsilon}(j)$. Let $x \in V_0^{\epsilon}(j)$ for some $j \geq 2$. Then we have $\tau_0^{\epsilon}(x) > j-1$, which implies

$$\sup\{t \geq 0 : |\det U(\Psi_0^t x)| < \epsilon\} > j-1.$$

Therefore,

$$\alpha\beta\left(e^{\alpha(j-1)}x_1 - e^{\beta(j-1)}x_2\right) < \epsilon$$

and, setting $y := e^{\alpha(j-1)}x_1$ and using that $x_1 > x_2^{\frac{\alpha}{\beta}}$,

$$\begin{aligned}
\frac{\epsilon}{\alpha\beta} &> e^{\alpha(j-1)}x_1 - e^{\beta(j-1)}x_2 \\
&> e^{\alpha(j-1)}x_1 - e^{\beta(j-1)}x_1^{\frac{\beta}{\alpha}} = y - y^{\frac{\beta}{\alpha}}.
\end{aligned} \tag{49}$$

As $y \mapsto y - y^{\frac{\beta}{\alpha}}$ is increasing for $y \geq 1$ and negative for $y \in (0,1)$, the definition of q implies $e^{\alpha(j-1)}x_1 < q$ and thus

$$x_1 < q e^{-\alpha(j-1)},$$

provided that $\epsilon < \alpha\beta$. Since $|\det U(x)| < \epsilon$ and thus $x_2 < x_1 + \frac{\epsilon}{\alpha\beta}$, the right side of the first line of (49) is also greater than

$$e^{\alpha(j-1)}x_1 - e^{\beta(j-1)}x_1 - e^{\beta(j-1)}\frac{\epsilon}{\alpha\beta},$$

whence it follows that

$$\begin{aligned}
x_1 < \frac{\epsilon}{\alpha\beta}\frac{1+e^{\beta(j-1)}}{e^{\alpha(j-1)}-e^{\beta(j-1)}} &= \frac{\epsilon}{\alpha\beta}\frac{1+e^{-\beta(j-1)}}{e^{(\alpha-\beta)(j-1)}-1} \\
&< \epsilon\frac{2}{\alpha\beta}\frac{1}{e^{(\alpha-\beta)(j-1)}-1} < \epsilon c' \frac{2}{\alpha\beta}e^{(\beta-\alpha)(j-1)}.
\end{aligned}$$

So far, we have shown that $x_1 < c^{-1}m^{\epsilon}(j)$. The asserted inequality $\tau_1^{\epsilon}(x) < m^{\epsilon}(j)$ then follows from $\tau_1^{\epsilon}(x) < cx_1$, which is equivalent to (48). It remains to consider the case $j=1$. Similarly to the case $j \geq 2$,

$$x_1 < \frac{2\epsilon}{\alpha\beta(e^{(\alpha-\beta)\epsilon^{\frac{\beta}{\alpha}}}-1)} \leq c^{-1}m^{\epsilon}(1).$$

We can then conclude as in the case $j \geq 2$. □

PROOF OF LEMMA 11: Fix $n \in \mathbb{N}$, $\lambda_0 > 0$, $\lambda_1 \in (-\alpha, 0]$ and $x \in \Gamma^\circ$. Since

$$\{x \in \Gamma^\circ : |\det U(x)| < \epsilon\} = \bigcup_{j=0}^{\infty} V_0^\epsilon(j) = \bigcup_{j=0}^{\infty} V_1^\epsilon(j),$$

and since the sets $(V_i^\epsilon(j))_{j\in\mathbb{N}_0}$ are disjoint for $i \in \{0,1\}$, we have

$$\widehat{\mathcal{M}}_n^{\epsilon,\lambda_0,\lambda_1}(x) = \sum_{(j_1,\dots,j_n)\in\mathbb{N}_0^n} \int_{t_n\in\mathbb{R}_+:\Psi_0^{t_n}(x)\in V_1^\epsilon(j_n)} dt_n\ e^{-\lambda_0 t_n} \int_{t_{n-1}\in\mathbb{R}_+:\Psi_0^{(t_{n-1},t_n)}(x)\in V_0^\epsilon(j_{n-1})} dt_{n-1}\ e^{-\lambda_1 t_{n-1}} \dots \int_{t_1\in\mathbb{R}_+:\Psi_0^{(t_1,\dots,t_n)}(x)\in V_{i_n}^\epsilon(j_1)} dt_1 e^{-\lambda_{1-i_n}t_1-\lambda_{i_n}\tau_{i_n}(\Psi_0^{(t_1,\dots,t_n)}x))}, \tag{50}$$

where $i_n = 0$ for n even, and $i_n = 1$ for n odd. Fix $(j_1,\dots,j_n) \in \mathbb{N}_0^n$ and a sequence of switching times $(t_2,\dots,t_n)$ such that

$$\Psi_0^{(t_k,\dots,t_n)}(x) \in V_{i_{n-k+1}}^\epsilon(j_k), \quad 2 \le k \le n.$$

We would like to estimate the integral in the third line of (50). Set $z := \Psi_0^{(t_2,\dots,t_n)}(x)$ and let $t_1 \in \mathbb{R}_+$ such that $\Psi_{i_{n+1}}^{t_1}(z) \in V_{i_n}^\epsilon(j_1)$. Since $z \in V_{i_{n-1}}^\epsilon(j_2)$, we have $|\det U(z)| < \epsilon$. Besides, as $\Psi_{i_{n+1}}^{t_1}(z) \in V_{i_n}^\epsilon(j_1)$, we have $\Psi_{i_{n+1}}^{t_1}(z) \in \Gamma^\circ$ and $|\det U(\Psi_{i_{n+1}}^{t_1}(z))| < \epsilon$. Hence,

$$t_1 < \tau_{i_{n+1}}^\epsilon(z) \le j_2 + 1. \tag{51}$$

For $j_2 \ge 1$, we even have $\tau_{i_{n+1}}^\epsilon(z) \le j_2$, but we work with this slightly worse estimate to avoid distinguishing between the cases $j_2 = 0$ and $j_2 \ge 1$. Recall that $\tau_i(x)$ was defined in Lemma 9. We claim that for every $\epsilon > 0$,

$$d(\epsilon) := \sup_{x\in\Gamma^\circ:|\det U(x)|<\epsilon} (\tau_0(x) - \tau_0^\epsilon(x)) < \infty. \tag{52}$$

Let $x \in \Gamma^\circ$ such that $|\det U(x)| < \epsilon$, and assume without loss of generality that $\tau_0(x)$ is strictly larger than $\tau_0^\epsilon(x)$. Then,

$$\tau_0(x) - \tau_0^\epsilon(x) = \tau_0(y(x)),$$

where

$$y(x) = \Psi_0^{\tau_0^\epsilon(x)}(x) \in \{z \in \Gamma^\circ : \det U(z) = \epsilon\}.$$

Since τ_0 is continuous and since $\tau_0(y)$ converges to 0 as y approaches either one of the endpoints of the line segment $\{z \in \Gamma^\circ : \det U(z) = \epsilon\}$, the claim in (52) follows. Together with (51), the fact that $\Psi_{i_{n+1}}^{t_1}(z) \in V_{i_n}^\epsilon(j_1)$, and the assumption $\lambda_0 > 0 \ge \lambda_1$, one obtains

$$e^{-\lambda_{1-i_n}t_1-\lambda_{i_n}\tau_{i_n}(\Psi_0^{(t_1,\dots,t_n)}x)} \le e^{-\lambda_1(j_1+j_2+d(\epsilon)+2)},$$

so the integral in the third line of (50) is bounded from above by

$$e^{-\lambda_1(j_1+j_2+d(\epsilon)+2)} \operatorname{Leb}\left(\left\{t_1 \in \mathbb{R}_+ : \Psi_{i_{n+1}}^{t_1}(z) \in V_{i_n}^\epsilon(j_1)\right\}\right),$$

where Leb denotes the Lebesgue measure. Fix a time

$$t_1 \in \left\{t_1 \in \mathbb{R}_+ : \Psi_{i_{n+1}}^{t_1}(z) \in V_{i_n}^\epsilon(j_1)\right\}$$

and define $y := \Psi^{t_1}_{i_{n+1}}(z)$. If $j_1 \geq 1$, Lemma 12 yields $\tau^\epsilon_{i_{n+1}}(y) < m^\epsilon(j_1)$, so

$$\text{Leb}\left(\left\{t_1 \in \mathbb{R}_+ : \Psi^{t_1}_{i_{n+1}}(z) \in V^\epsilon_{i_n}(j_1)\right\}\right) \leq m^\epsilon(j_1). \tag{53}$$

If $j_1 = 0$, the expression on the lefthand side of (53) is bounded from above by

$$\sup_{y \in V^\epsilon_{i_{n-1}}(j_2)} \tau^\epsilon_{i_{n+1}}(y) = j_2 + \epsilon^{\frac{\beta}{\alpha}} \mathbb{1}_{j_2=0}.$$

For $i, j \in \mathbb{N}_0$, define

$$h(i,j) = e^{-\lambda_1(j+1)} \left(m^\epsilon(i)\mathbb{1}_{i\geq 1} + (j + \epsilon^{\frac{\beta}{\alpha}} \mathbb{1}_{j=0})\mathbb{1}_{i=0}\right).$$

The integral in the third line of (50) is thus less than

$$e^{-\lambda_1(j_1+1+d(\epsilon))} h(j_1, j_2).$$

A similar estimate without the factor $e^{-\lambda_1(j_1+1+d(\epsilon))}$ applies to each of the other integrals in (50), with the crucial exception of

$$\int_{t_n \in \mathbb{R}_+ : \Psi^{t_n}_0(x) \in V^\epsilon_1(j_n)} e^{-\lambda_0 t_n}\, dt_n \leq \frac{1}{\lambda_0}. \tag{54}$$

Then,

$$\begin{aligned} &\widehat{\mathcal{M}}^{\epsilon,\lambda_0,\lambda_1}_n(x) \\ \leq &\frac{e^{-\lambda_1 d(\epsilon)}}{\lambda_0} \sum_{(j_1,\ldots,j_n)\in\mathbb{N}_0^n} e^{-\lambda_1(j_1+1)} \prod_{l=1}^{n-1} h(j_l, j_{l+1}) = \frac{e^{-\lambda_1 d(\epsilon)}}{\lambda_0} \sum_{A\subset\{1,\ldots,n\}} J_A, \end{aligned} \tag{55}$$

where

$$J_A = \sum_{(j_1,\ldots,j_n)\in\mathbb{N}_A} e^{-\lambda_1(j_1+1)} \prod_{l=1}^{n-1} h(j_l, j_{l+1})$$

and

$$\mathbb{N}_A = \{(j_1,\ldots,j_n) \in \mathbb{N}_0^n : j_l > 0 \text{ iff } l \in A\}, \quad A \subset \{1,\ldots,n\}.$$

Fix a set $A \subset \{1,\ldots,n\}$. We want to estimate the term J_A. If $A = \emptyset$,

$$J_A = e^{-\lambda_1} h(0,0)^{n-1} = \epsilon^{-\frac{\beta}{\alpha}} \left(\epsilon^{\frac{\beta}{\alpha}} e^{-\lambda_1}\right)^n. \tag{56}$$

Now assume $A \neq \emptyset$. We call a subset B of A a *connected component* if $k < l < m$ for $k, m \in B$ and $l \in \{1,\ldots,n\}$ implies $l \in B$ and if no subset of A that strictly contains B has this property. The set A can be written as the disjoint union of its connected components, and the number of connected components of the set A ranges from 1 to the cardinality of A. We call the number of indices between two adjacent connected components B_1 and B_2 the *gap* between B_1 and B_2.

Let $B_1,\ldots,B_m$ be the connected components of a nonempty set $A \subset \{1,\ldots,n\}$, written in increasing order, i.e. B_1 contains the smallest index in A, B_2 is adjacent to B_1, etc. We denote the sizes of $B_1,\ldots,B_m$ by $s_1,\ldots,s_m$, respectively, and write $B_l = \{j^l_1,\ldots,j^l_{s_l}\}$ for $1 \leq l \leq m$. For $1 \leq l \leq m-1$, let g_l denote the gap between B_l and B_{l+1}. Let g_0 be the number of indices preceding B_1 and let g_m be the number of indices following B_m. Then, $\vartheta := (g_0-1)^+ + \sum_{l=1}^{m-1}(g_l-1) + (g_m-1)^+$ is the number of instances in which two subsequent indices are both not in A. Depending on whether $g_0 > 0$ or $g_m > 0$, there are four cases to consider. We only present the case $g_0, g_m > 0$, and thus omit the cases where $g_0 = 0$ or $g_m = 0$. We have

$$J_A = e^{-\lambda_1} h(0,0)^{\vartheta} \prod_{l=1}^{m} \Bigg(\sum_{j_1^l=1}^{\infty} h(0, j_1^l) \sum_{j_2^l=1}^{\infty} h(j_1^l, j_2^l) \dots \sum_{j_{s_l}^l=1}^{\infty} h(j_{s_l-1}^l, j_{s_l}^l) h(j_{s_l}^l, 0) \Bigg). \tag{57}$$

For $j_1^l, \dots, j_{s_l}^l \geq 1$,

$$h(0, j_1^l) h(j_1^l, j_2^l) \dots h(j_{s_l-1}^l, j_{s_l}^l) h(j_{s_l}^l, 0) = e^{-\lambda_1 j_1^l} \prod_{i=1}^{s_l} e^{-\lambda_1 (j_i^l + 1)} m^{\epsilon}(j_i^l).$$

Thus, the right side of (57) can be written as

$$e^{-\lambda_1} \left(\epsilon^{\frac{\beta}{\alpha}} \, e^{-\lambda_1} \right)^{\vartheta} e^{-\lambda_1 m} \prod_{l=1}^{m} \Bigg(e^{-\lambda_1 s_l} \left(\sum_{j_1^l=1}^{\infty} j_1^l e^{-\lambda_1 j_1^l} m^{\epsilon}(j_1^l) \right) \left(\sum_{j=1}^{\infty} e^{-\lambda_1 j} m^{\epsilon}(j) \right)^{s_l - 1} \Bigg).$$

It follows that

$$J_A \leq e^{-\lambda_1} \left(\epsilon^{\frac{\beta}{\alpha}} \, e^{-\lambda_1} \right)^{\vartheta} e^{-\lambda_1 (m + |A|)} b(\epsilon)^{|A|}, \tag{58}$$

where

$$b(\epsilon) = \sum_{j=1}^{\infty} j e^{-\lambda_1 j} m^{\epsilon}(j),$$

and where $|A| = \sum_{l=1}^{m} s_l$ is the cardinality of A. Since $\lambda_1 \in (-\alpha, 0]$, Lemma 12 implies that $\lim_{\epsilon \downarrow 0} b(\epsilon) = 0$. If we set

$$\tilde{f}(\epsilon) = \left(\epsilon^{\frac{\beta}{\alpha}} \, e^{-\lambda_1} \right) \vee b(\epsilon),$$

we obtain

$$J_A \leq e^{-\lambda_1} \left(e^{-2\lambda_1} \tilde{f}(\epsilon)^{\frac{1}{3}} \right)^{n} \tag{59}$$

because $\vartheta + |A| \geq \frac{n-1}{2}$ and $m + |A| \leq 2n$. This completes the estimate of J_A in the case $g_0, g_n > 0$. In each of the remaining three cases, one can show without much effort that J_A is less than or equal to the expression on the right side of (58).

From (55), (56) and (59), we infer that

$$\widehat{\mathcal{M}}_n^{\epsilon, \lambda_0, \lambda_1}(x) \leq c(\epsilon) f(\epsilon)^n$$

for

$$c(\epsilon) = \frac{e^{-\lambda_1 d(\epsilon)}}{\lambda_0} \epsilon^{-\frac{\beta}{\alpha}} e^{-\lambda_1}, \quad f(\epsilon) = 2e^{-2\lambda_1} \left(\epsilon^{\frac{\beta}{\alpha}} \vee \tilde{f}(\epsilon)^{\frac{1}{3}} \right).$$

This completes the proof of Lemma 11 and thus of Theorem 2, part (1). ☐

8.4. **Proof of Theorem 2, part (2).** The proof of part (2) is quite similar to the proof of part (1) for $\lambda_1 \in (\beta, \alpha+\beta]$. Let K be a compact subset of Γ_0, and recall that $\widehat{M}$ was defined in (45). Following the proof of part (1), all we need to show is the following lemma.

Lemma 13. *There are functions $c(\epsilon)$ and $f(\epsilon)$ such that $\lim_{\epsilon\downarrow 0} f(\epsilon) = 0$ and*

$$\widehat{\mathcal{M}}_n^{\epsilon,\lambda_0,\lambda_1}(x) \le c(\epsilon) f(\epsilon)^n, \quad n \in \mathbb{N},\ x \in K \cap \Gamma^\circ,\ \lambda_0, \lambda_1 \in (-\alpha, 0]. \tag{60}$$

Proof: Fix $n \in \mathbb{N}$, $\lambda_0, \lambda_1 \in (-\alpha, 0]$ and $x \in K \cap \Gamma^\circ$. As in the proof of Lemma 11, we use the representation for $\widehat{\mathcal{M}}_n^{\epsilon,\lambda_0,\lambda_1}(x)$ in (50). If we let $\lambda = \lambda_0 \wedge \lambda_1$, the integral in the third line of (50) is less than

$$e^{-\lambda(j_1+1+d(\epsilon))} h(j_1, j_2),$$

where

$$h(i,j) = e^{-\lambda(j+1)} \left(m^\epsilon(i) \mathbb{1}_{i\ge 1} + \left(j + \epsilon^{\frac{\beta}{\alpha}} \mathbb{1}_{j=0} \right) \mathbb{1}_{i=0} \right).$$

As in the proof of Lemma 11, there are similar estimates for the other integrals in (50), but now the integral on the left side of (54) is less than

$$h_K(j_n) := e^{-\lambda(T_K+1)} \left(m^\epsilon(j_n) \mathbb{1}_{j_n \ge 1} + T_K \mathbb{1}_{j_n=0} \right).$$

Here, T_K is a positive integer that depends only on the compact set K, and whose existence follows from the assumption $(0,0) \notin K$. The proof can then essentially be completed as the one of Lemma 11, with $\frac{1}{\lambda_0}$ replaced by $h_K(j_n)$. □

8.5. **Proof of Theorem 2, part (3).** The proof strategy is similar to the one from the proof of part (1) for $\lambda_1 > \alpha + \beta$. Here, however, we need to make sure that the compositions of backward trajectories do not become arbitrarily close to the critical points of u_0 and u_1. This is accomplished by letting the width of the strip around the diagonal shrink to zero as we move backward in time. The procedure only works because we require K to be a positive distance away from the boundary curve $\partial\Gamma_0$. For $n \in \mathbb{N}$, $x \in \Gamma^\circ$, and $\epsilon > 0$, let

$$\sigma M_n^\epsilon(x) = \left\{ \mathbf{t} \in T_0^n(x) : \left| \det U(\Psi_0^{(t_{n-j},\dots,t_n)} x) \right| < \epsilon 2^{-j},\ 0 \le j \le n-1 \right\},$$
$$\sigma S_n^\epsilon(x) = \left\{ (t_1, t_2, \mathbf{t}) \in T_0^{n+1}(x) : \mathbf{t} \in \sigma M_{n-1}^\epsilon(x),\ \left| \det U(\Psi_0^{(t_2,\mathbf{t})} x) \right| > \epsilon 2^{-(n-1)} \right\}.$$

For $h \in L^1(\Gamma^\circ)$, define

$$\sigma\mathcal{A}_n^\epsilon h(x) = \int_{\sigma S_n^\epsilon(x)} \lambda_0(n+1) e^{-\langle \lambda_0^{(n+1)}, \mathbf{t}\rangle} J_0^{\mathbf{t}}(x) h(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t},$$
$$\sigma\mathcal{B}_n^\epsilon h(x) = \int_{\sigma M_n^\epsilon(x)} \lambda_0(n) e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} J_0^{\mathbf{t}}(x) h(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t}.$$

The letter σ stands for “shrinking” (referring to the strip around the diagonal) and is meant to help distinguish the notation from the one introduced at the beginning of Subsection 8.1.

We need analogs of Lemmas 6 and 8, which we state in Lemma 14 below. The proofs of these modified statements are almost identical to the proofs of the original ones, and we omit them.

Lemma 14. *For any $x \in \Gamma^\circ$ and $\epsilon > 0$, the following statements hold.*

(1) *For any* $n \in \mathbb{N}$,

$$\rho_0(x) = \sum_{k=1}^{n} \sigma\mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) + \sigma\mathcal{B}_n^\epsilon \rho_{i_n}(x) = \sum_{k=1}^{\infty} \sigma\mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x).$$

(2) *We have*

$$\sigma\mathcal{A}_1^\epsilon \rho_{i_2}(x) = \mathcal{I}_0^\epsilon(x),$$

and for any $k > 1$,

$$\sigma\mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) = \int_{\sigma M_{k-1}^\epsilon(x)} \lambda_0(k-1) e^{\langle(\alpha+\beta)\mathbb{1} - \lambda_0^{(k-1)}, \mathbf{t}\rangle} \mathcal{I}_{i_{k+1}}^{\epsilon 2^{-(k-1)}}(\Psi_0^{\mathbf{t}} x)\ d\mathbf{t}.$$

Next, we formulate an analog of Lemma 10. For $n \in \mathbb{N}$, $\epsilon > 0$, $x \in \Gamma^\circ$, and $\lambda_0, \lambda_1 \in \mathbb{R}$, set

$$\sigma\mathcal{M}_n^{\epsilon,\lambda_0,\lambda_1}(x) = \int_{\sigma M_n^\epsilon(x)} e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle}\ d\mathbf{t}.$$

Lemma 15. *For any* $\lambda_0, \lambda_1 \in \mathbb{R}$*, there is a function* $f(\epsilon)$ *such that* $\lim_{\epsilon\downarrow 0} f(\epsilon) = 0$ *and*

$$\sigma\mathcal{M}_{n+1}^{\epsilon,\lambda_0,\lambda_1}(x) \le f(\epsilon)\sigma\mathcal{M}_n^{\epsilon,\lambda_0,\lambda_1}(x), \quad n \in \mathbb{N},\ x \in K \cap \Gamma^\circ.$$

Before proving Lemma 15, we carry out some preliminary work. For $\delta \in (0, \frac{1}{2})$, let

$$\Gamma(\delta) = \{x \in \Gamma^\circ : \delta < x_2 < 1 - \delta\}.$$

Lemma 16. *There is* $\delta > 0$ *such that for* $\epsilon > 0$ *sufficiently small,*

$$\Psi_0^{\mathbf{t}}(x) \in \Gamma(\delta), \quad x \in K \cap \Gamma^\circ,\ n \in \mathbb{N},\ \mathbf{t} \in \sigma M_n^\epsilon(x).$$

Proof: We will at times use the notation $[x]_2$ for the second component of a point $x \in \mathbb{R}^2$. For $\delta > 0$ and $n \in \mathbb{N}$, we set

$$\mathcal{S}(\delta, n) := 2\delta - \frac{\delta}{2}\sum_{k=1}^{n} 2^{-k+1}$$

to simplify notation. Due to symmetries and an induction argument, it is enough to show that there is $\delta > 0$ such that for $\epsilon > 0$ sufficiently small, the following statements hold:

(1) For any $x \in K \cap \Gamma^\circ$ and $t \in \sigma M_1^\epsilon(x)$, one has

$$\Psi_0^t(x) \in \Gamma(2\delta);$$

(2) For any even positive integer n, $y \in \Gamma(\mathcal{S}(\delta, n))$ such that $|\det U(y)| < \epsilon 2^{-(n-1)}$, and $t > 0$ such that $\Psi_0^t(y) \in \Gamma^\circ$ and $|\det U(\Psi_0^t y)| < \epsilon 2^{-n}$, one has

$$\Psi_0^t(y) \in \Gamma(\mathcal{S}(\delta, n+1)).$$

First we specify δ. For $x \in \mathbb{R}^2$ and $t \in \mathbb{R}$, consider the function

$$\delta(x) = x_2 \vee (1 - x_2).$$

According to Lemma 5, for any $i \in \{0, 1\}$ and $x \in \Gamma^\circ$, there is a unique $\theta_i(x) \in \mathbb{R}$ such that $\det U(\Psi_i^{\theta_i(x)} x) = 0$ and $\Psi_i^{\theta_i(x)}(x) \in \Gamma^\circ$. We define

$$\zeta = \sup_{x \in K \cap \Gamma^\circ} \delta\left(\Psi_0^{\theta_0(x)} x\right),$$

which is strictly less than 1 because K is compact and does not intersect $\partial\Gamma_0$. Then, we set

$$\delta := \frac{1-\zeta}{3} \wedge \frac{1}{2}\inf_{x\in K} x_2.$$

Let $\tilde{\epsilon} > 0$ be so small that the closure of

$$\Xi(\delta,\tilde{\epsilon}) := \{x \in \Gamma_\delta : |\det U(x)| < \tilde{\epsilon}\}$$

is contained in Γ°. Set

$$\vartheta = \inf_{x\in(K\cap\Gamma^\circ)\cup\Xi(\delta,\tilde{\epsilon})} [\Psi_0^{\theta_0(x)}(x)]_1 > 0.$$

For any $x \in \mathbb{R}^2$ such that $\delta \le x_1 = x_2 \le 1$, the formula in part (2) of Lemma 5 implies

$$\frac{d}{ds}\det U(\Psi_0^s x)|_{s=0} \ge \alpha\beta(\alpha-\beta)\delta,$$

so by a compactness argument there is $r > 0$ such that

$$\frac{d}{ds}\det U(\Psi_0^s x)|_{s=0} \ge \frac{\alpha\beta}{2}(\alpha-\beta)\delta \tag{61}$$

for every $x \in \mathbb{R}^2$ such that $\delta \le x_2 \le 1$ and $|\det U(x)| < r$. Next, observe that since u_0 and u_1 are bounded on the compact set Γ, there is $C > 0$ such that

$$|\partial_t \Psi_i^t(x)| = |u_i(\Psi_i^t(x))| \le C$$

for every $i \in \{0,1\}$, $x \in \Gamma^\circ$, and $t \ge 0$ for which $\Psi_i^t(x) \in \Gamma^\circ$.

We proceed to the proof of statements (1) and (2). We will assume that ϵ is sufficiently small with respect to δ, $\tilde{\epsilon}$, ϑ, and C for the estimates given above to hold. First we prove statement (1), which plays the role of the base case in an induction argument. Fix a point $x \in K \cap \Gamma^\circ$. For any $t \in \sigma M_1^\epsilon(x)$, one has

$$2\delta \le x_2 < [\Psi_0^t(x)]_2,$$

because $t \mapsto [\Psi_0^t(x)]_2$ is increasing. To show the estimate

$$1 - 2\delta > [\Psi_0^t(x)]_2,$$

assume without loss of generality that $t > \theta_0(x)$. Set $c(s) = (c_1(s), c_2(s)) := \Psi_0^s(x)$ for $s \in [\theta_0(x), t]$. Then,

$$\epsilon > \det U(c(t)) = \det U(c(t)) - \det U(c(\theta_0(x))) = \int_{\theta_0(x)}^t \frac{d}{ds}\det U(c(s))\, ds. \tag{62}$$

For $s \in [\theta_0(x), t]$, one has, again by part (2) of Lemma 5,

$$\frac{d}{ds}\det U(c(s)) \ge \alpha\beta(\alpha-\beta)c_1(\theta_0(x)) \ge \alpha\beta(\alpha-\beta)\vartheta.$$

Together with (62), this yields

$$t - \theta_0(x) < \frac{\epsilon}{\alpha\beta(\alpha-\beta)\vartheta}.$$

By the mean-value theorem, there is $s^* \in (\theta_0(x), t)$ such that

$$c_2(t) - c_2(\theta_0(x)) = (t - \theta_0(x))c_2'(s^*).$$

Then,

$$\begin{aligned}c_2(t) =&c_2(\theta_0(x)) + c_2(t) - c_2(\theta_0(x))\\ \le&\zeta + (t-\theta_0(x))c_2'(s^*) \le \zeta + \frac{\epsilon C}{\alpha\beta(\alpha-\beta)\vartheta} \le 1-3\delta + \frac{\epsilon C}{\alpha\beta(\alpha-\beta)\vartheta} < 1-2\delta,\end{aligned}$$

which completes the proof of statement (1).

We proceed to the proof of statement (2). Let n be an even positive integer, let $y \in \Gamma(\mathcal{S}(\delta,n))$ such that $|\det U(y)| < \epsilon 2^{-(n-1)}$, and let $t > 0$ such that $\Psi_0^t(y) \in \Gamma^\circ$ and $|\det U(\Psi_0^t y)| < \epsilon 2^{-n}$. Then,

$$[\Psi_0^t(y)]_2 > y_2 > \mathcal{S}(\delta,n) > \mathcal{S}(\delta,n+1).$$

It remains to show

$$[\Psi_0^t(y)]_2 < 1 - \mathcal{S}(\delta,n+1). \tag{63}$$

As in the proof of statement (1), there is no loss of generality in assuming $t > \theta_0(y)$. Set $d(s) = (d_1(s), d_2(s)) := \Psi_0^s(y)$ for $s \in [\theta_0(y), t]$. Since $y \in \Xi(\delta, \tilde{\epsilon})$, one has $d_1(\theta_0(y)) \ge \vartheta$. Thus, we can essentially repeat the argument from the proof of statement (1) to obtain

$$t - \theta_0(y) < \frac{\epsilon 2^{-n}}{\alpha\beta(\alpha-\beta)\vartheta}$$

and

$$d_2(t) \le d_2(\theta_0(y)) + \frac{\epsilon 2^{-n} C}{\alpha\beta(\alpha-\beta)\vartheta}.$$

The next step consists in estimating $d_2(\theta_0(y))$ from above. Assume without loss of generality that $d_2(\theta_0(y)) > y_2$. Then,

$$\epsilon 2^{-(n-1)} > \int_0^{\theta_0(y)} \frac{d}{ds} \det U(d(s))\, ds \tag{64}$$

in analogy to (62). Next, we estimate $\frac{d}{ds}\det U(d(s))$ from below. For fixed $s \in [0, \theta_0(y)]$, we claim that

$$\det U(d(s)) > -r,$$

where one should recall that r was introduced in relation to (61). Suppose the claim doesn't hold. Since $\det U(d(0)) > -r$, there is $s^* \in (0, s]$ such that $\det U(d(s^*)) = -r$ and $\det U(d(t)) > -r$ for every $t \in [0, s^*)$. Then,

$$\begin{aligned}-r =&\det U(d(s^*)) = \det U(y) + \det U(d(s^*)) - \det U(d(0))\\ >& -\epsilon 2^{-(n-1)} + \int_0^{s^*} \frac{d}{ds}\det U(d(s))\, ds \ge -\epsilon + s^* \frac{\alpha\beta}{2}(\alpha-\beta)\delta > -r,\end{aligned}$$

a contradiction. As a result, the integral on the righthand side of (64) is bounded from below by

$$\theta_0(y)\frac{\alpha\beta}{2}(\alpha-\beta)\delta.$$

Hence,

$$\theta_0(y) < \frac{\epsilon 2^{-(n-2)}}{\alpha\beta(\alpha-\beta)\delta},$$

and we obtain the estimate

$$d_2(\theta_0(y)) = y_2 + d_2(\theta_0(y)) - d_2(0) < 1 - \mathcal{S}(\delta,n) + \frac{\epsilon 2^{-(n-2)}}{\alpha\beta(\alpha-\beta)\delta}C.$$

This yields

$$d_2(t) < 1 - \mathcal{S}(\delta, n) + \left(\frac{4}{\delta} + \frac{1}{\vartheta}\right) \frac{\epsilon 2^{-n} C}{\alpha\beta(\alpha-\beta)} \leq 1 - \mathcal{S}(\delta, n+1).$$

□

Proof of Lemma 15: Let $n \in \mathbb{N}$, $x \in K \cap \Gamma^\circ$, $\epsilon > 0$, and $\lambda_0, \lambda_1 \in \mathbb{R}$. Let $\lambda = |\lambda_0| \vee |\lambda_1|$. As an immediate consequence of Lemma 16, there is $\delta > 0$, independent of n and x, such that for $\epsilon > 0$ sufficiently small,

$$\Psi_0^{(t_{n+1-j},\ldots,t_{n+1})}(x) \in \Gamma(\delta), \quad 0 \leq j \leq n, \ \mathbf{t} \in \sigma M_{n+1}^\epsilon(x). \tag{65}$$

Let $\mathbf{t} = (t_1, t_2, \ldots, t_{n+1}) \in \sigma M_{n+1}^\epsilon(x)$ and set $y := \Psi_0^{(t_2,\ldots,t_{n+1})}(x)$. As we saw in the proof of Lemma 16,

$$t_1 = t_1 - \theta_{i_n}(y) + \theta_{i_n}(y) < \frac{\epsilon 2^{-n}}{\alpha\beta(\alpha-\beta)\vartheta} + \frac{\epsilon 2^{-(n-2)}}{\alpha\beta(\alpha-\beta)\delta} \leq c\epsilon,$$

where $c > 0$ is a constant that does not depend on n. Hence,

$$\sigma\mathcal{M}_{n+1}^{\epsilon,\lambda_0,\lambda_1}(x) \leq \int_{\sigma M_n^\epsilon(x)} d\mathbf{t}\ e^{-\langle \lambda_0^{(n)}, \mathbf{t}\rangle} \int_0^{c\epsilon} dt_1\ e^{\lambda t_1} \leq c\epsilon e^{\lambda c\epsilon} \sigma\mathcal{M}_n^{\epsilon,\lambda_0,\lambda_1}(x).$$

As $\lim_{\epsilon\downarrow 0} c\epsilon e^{\lambda c \epsilon} = 0$, this completes the proof. □

Proof of Theorem 2, part (3): By Lemma 14, we need to show that

$$\sup_{x \in K\cap\Gamma^\circ} \sum_{k=1}^\infty \sigma\mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) < \infty.$$

Again by Lemma 14, we have for $x \in K \cap \Gamma^\circ$

$$\sum_{k=1}^\infty \sigma\mathcal{A}_k^\epsilon \rho_{i_{k+1}}(x) = \mathcal{I}_0^\epsilon(x) + \sum_{k=2}^\infty \int_{\sigma M_{k-1}^\epsilon(x)} \lambda_0(k-1) e^{\langle(\alpha+\beta)\mathbb{1} - \lambda_0^{(k-1)}, \mathbf{t}\rangle} \mathcal{I}_{i_{k+1}}^{\epsilon 2^{-(k-1)}}(\Psi_0^{\mathbf{t}} x)\, d\mathbf{t}.$$

By Lemma 9, the righthand side is less than

$$\begin{aligned} &\frac{2\lambda_0\lambda_1}{\epsilon} e^{(\alpha+\beta-\lambda_0)\tau_0(x)} \\ &+ \sum_{k=2}^\infty \frac{2\lambda_0\lambda_1}{\epsilon} 2^{k-1} \lambda_0(k-1) \int_{\sigma M_{k-1}^\epsilon(x)} e^{\langle(\alpha+\beta)\mathbb{1} - \lambda_0^{(k)}, (\tau_{i_{k+1}}(\Psi_0^{\mathbf{t}} x), \mathbf{t})\rangle}\ d\mathbf{t}. \end{aligned} \tag{66}$$

For any $k \geq 2$ and $\mathbf{t} \in \sigma M_{k-1}^\epsilon(x)$, we have $|\det U(\Psi_0^{\mathbf{t}} x)| < \epsilon$ and, on account of (65), we also have $\Psi_0^{\mathbf{t}}(x) \in \Gamma(\delta)$. Therefore,

$$\tau_{i_{k+1}}(\Psi_0^{\mathbf{t}} x) \leq c,$$

where c is a finite constant that does not depend on ϵ. The expression in the second line of (66) is thus bounded from above by

$$e^{(\alpha+\beta)c} \frac{2\lambda_0\lambda_1}{\epsilon} \sum_{k=2}^\infty 2^{k-1} \lambda_0(k-1) \sigma\mathcal{M}_{k-1}^{\epsilon,\lambda_0-(\alpha+\beta),\lambda_1-(\alpha+\beta)}(x). \tag{67}$$

By Lemma 15, for $k \geq 2$,

$$\sigma\mathcal{M}_{k-1}^{\epsilon,\lambda_0-(\alpha+\beta),\lambda_1-(\alpha+\beta)}(x) \leq f(\epsilon)^{k-2} \sigma\mathcal{M}_1^{\epsilon,\lambda_0-(\alpha+\beta),\lambda_1-(\alpha+\beta)}(x) \leq f(\epsilon)^{k-2} \hat{c} e^{(\alpha+\beta)\hat{c}},$$

where $\lim_{\epsilon\downarrow 0} f(\epsilon) = 0$ and

$$\hat{c} = \sup_{x\in K\cap\Gamma^\circ} \sup\{t \geq 0 : \det U(\Psi_0^t x) \leq \epsilon\}.$$

Hence, the expression in (67) is bounded from above by

$$e^{(\alpha+\beta)(c+\hat{c})} \frac{2\lambda_0\lambda_1\hat{c}}{\epsilon f(\epsilon)} \sum_{k=1}^{\infty} (2f(\epsilon))^k \lambda_0(k),$$

which doesn't depend on x and is finite for ϵ sufficiently small. □

Acknowledgments: JCM thanks the NSF grant DMS-1613337 for partial support of this work. JCM and SDL thank Michael Reed for many useful discussions earlier in their work in randomly switched ODEs and this example in particular. TH gratefully acknowledges support through SNF grant $200021 - 175728/1$. He thanks Florent Malrieu for helpful suggestions related to this example, and Jean-Baptiste Bardet, Michel Benaïm, Florent Malrieu, Edouard Strickler, and Pierre-André Zitt for useful discussions about invariant densities for this and other two-dimensional examples. YB is grateful to NSF for partial support of his work via grants DMS-1460595 and DMS-1811444. SDL thanks the NSF grants DMS-1944574 and DMS-1814832 for partial support of this work. Finally, we would like to thank three anonymous referees whose comments and suggestions helped to improve the paper.

References

[1] Yuri Bakhtin and Tobias Hurth. Invariant densities for dynamical systems with random switching. *Nonlinearity*, 25(10):2937–2952, 2012.

[2] Yuri Bakhtin, Tobias Hurth, Sean D Lawley, and Jonathan C Mattingly. Smooth invariant densities for random switching on the torus. *Nonlinearity*, 31(4):1331, 2018.

[3] Yuri Bakhtin, Tobias Hurth, and Jonathan C. Mattingly. Regularity of invariant densities for 1d-systems with random switching. *Nonlinearity*, 28:3755–3787, 2015.

[4] 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.

[5] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. Qualitative properties of certain piecewise deterministic Markov processes. *Ann. Inst. Henri Poincaré Probab. Stat.*, 51(3):1040–1075, 2015.

[6] P C Bressloff. Stochastic switching in biology: from genotype to phenotype. *Journal of Physics A: Mathematical and Theoretical*, 50(13):133001, 2017.

[7] P C Bressloff and S D Lawley. Dynamically active compartments coupled by a stochastically-gated gap junction. *J. Nonlinear Sci*, 2017.

[8] M. H. A. Davis. *Markov models and optimization*, volume 49 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1993.

[9] A. Faggionato, D. Gabrielli, and M. Ribezzi Crivellari. Non-equilibrium thermodynamics of piecewise deterministic Markov processes. *J. Stat. Phys.*, 137(2):259–304, 2009.

[10] Reuben Hersh. The birth of random evolutions. *Math. Intelligencer*, 25(1):53–60, 2003.

[11] C Jia, M Q Zhang, and H Qian. Emergent lévy behavior in single-cell stochastic gene expression. *Phys Rev E*, 96(4):040402, 2017.

[12] Thomas B Kepler and Timothy C Elston. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophysical journal*, 81(6):3116–3136, 2001.

[13] S D Lawley. Boundary value problems for statistics of diffusion in a randomly switching environment: PDE and SDE perspectives. *SIAM J Appl Dyn Syst*, 15, 2016.

[14] S D Lawley. A probabilistic analysis of volume transmission in the brain. *SIAM J Appl Math*, 78(2):942–962, 2018.

[15] Sean D. Lawley, Jonathan C. Mattingly, and Michael C. Reed. Stochastic switching in infinite dimensions with applications to random parabolic PDE. *SIAM J. Math. Anal.*, 47(4):3035–3063, 2015.

[16] Florent Malrieu. Some simple but challenging Markov processes. *Ann. Fac. Sci. Toulouse Math. (6)*, 24(4):857–883, 2015.

[17] J Paulsson. Models of stochastic gene expression. *Physics of life reviews*, 2(2):157–175, 2005.

[18] Michael W Smiley and Stephen R Proulx. Gene expression dynamics in randomly varying environments. *Journal of mathematical biology*, 61(2):231–251, 2010.

[19] G. George Yin and Chao Zhu. *Hybrid switching diffusions*, volume 63 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2010. Properties and applications.

Courant Institute of Mathematical Sciences, New York University, 251 Mercer St, New York, NY 10012 USA

Université de Neuchâtel, Institut de mathématiques, Rue Emile-Argand 11, CH-2000 Neuchâtel

Department of Mathematics, University of Utah, Salt Lake City, UT 84112 USA

Mathematics Department and Department of Statistical Science, Duke University, Durham, NC 27708 USA