Geometric Ergodicity for Piecewise Contracting Processes with Applications for Tropical Stochastic Lattice Models

ANDREW J MAJDA
Courant Institute

AND

XIN T TONG
Courant Institute

Abstract

Stochastic lattice models are increasingly prominent as a way to capture highly intermittent unresolved features of moist tropical convection in climate science and as continuum mesoscopic models in material science. Stochastic lattice models consist of suitably discretized continuum partial differential equations interacting with Markov jump processes at each lattice site with transition rates depending on the local value of the continuum equation; they are a special case of piecewise deterministic Markov processes but often have an infinite state space and unbounded transition rates. Here a general theorem on geometric ergodicity for piecewise deterministic contracting processes is developed with full generality to apply to stochastic lattice models. A highly nontrivial application to the stochastic skeleton model for the Madden-Julian oscillation (Thual et al. 2013) is developed here where there is an infinite state space with unbounded and also degenerate transition rates. Geometric ergodicity for the stochastic skeleton model guarantees exponential convergence to a unique invariant measure which defines the statistical tropical climate of the the model. Another application of the general framework is developed here for stochastic lattice models designed to capture intermittent fluctuation in the simplest tropical climate models. Other straightforward applications to models motivated by material science are mentioned briefly here. © 2000 Wiley Periodicals, Inc.

1 Introduction

Stochastic lattice models consist of suitably discretized continuum partial differential equations (PDE) interacting with Markov jump processes at each lattice site with transition rates depending on the local value of the continuum equations. Stochastic lattice models are increasingly prominent as a way to capture highly intermittent unresolved features of moist tropical convection in climate science [31, 28, 30, 27, 13, 14, 11, 26] and as

Communications on Pure and Applied Mathematics, Vol. 000, 0001–0042 (2000)

continuum mesoscopic models in material science [21, 22, 23, 24]. Such models are also likely to be useful in neural science and economics among other applications in the near future. Stochastic lattice models are a special case of piecewise deterministic Markov process (PDMP) [9, 10, 20], but often have an infinite state space with unbounded transition rates and other degeneracy.

Geometric ergodicity guarantees the exponential convergence to a unique statistical invariant measure and is an important step in the mathematical analysis of stochastic dynamical systems [37, 36, 16, 3, 7]. Here we build on these earlier works and formulate a new abstract theorem on geometric ergodicity for contracting PDMP which allows for an infinite state space and suitable unbounded transition rates. This abstract theorem is used in a highly nontrivial example, to prove the geometric ergodicity of the stochastic skeleton model of the Madden-Julian oscillation (MJO) [40, 39], where there is an initinite state space with unbounded and also degenerate transitions.

The outline of the present paper is as follows. In Section 2 we formulate and present two stochastic lattice models for moist tropical convection in climate science as motivation for further developments in the paper. One model is the stochastic skeleton model for the MJO [40, 39] with small nonzero damping; the second model is a stochastic lattice model designed to capture intermittent fluctuations [31, 28, 30] in the simplest tropical climate models [15, 32, 29]. With this background, the general theorem on contracting PDMP's is formulated and proved in Section 3 and 4 using Lyapunov functions and perturbation analysis. Application to stochastic parameterization for the simplest stochastic climate model is presented in Section 5. The highly nontrivial verification of geometric ergodicity for the stochastic skeleton model utilizing both the general theorem and detailed special structure of the stochastic skeleton model is presented in Section 6. Details of crucial Lyapunov stability bounds for these models are presented in the appendix. The theorem from Section 3 also applies in a simple straight forward fashion to stochastic lattice models motivated by material science [23, 24], but this is left as an exercise for the interested readers. We end the introduction with a brief heuristic discussion of the main theorem presented in Section 3 and 4.

Stochastic lattice models consist of an ordinary differential equation (ODE) system $\mathbf{U}_t$ which represents a discretized continuum PDE and a Markov jump process η_t, which represents random fluctuations. Their interaction can be parameterized by a vector field ψ that drives $\mathbf{U}_t$, and the jump intensity λ that governs η_t. In other words, the joint dynamics can be concisely described as:

$$d\mathbf{U}_t = \psi(\mathbf{U}_t, \eta_t)dt, \quad \mathbb{P}(\eta_{t+\Delta t} = \tilde{\eta} | \eta_t, \mathbf{U}_t) = \lambda(\mathbf{U}_t, \eta_t, \tilde{\eta})\Delta t.$$

The formal infinitesimal generator $\mathscr{L}$ of this process is given by:

$$\mathscr{L}f(u,\eta) = \psi(u,\eta)\nabla_u f(u,\eta) + \sum_{\tilde{\eta}} \lambda(u,\eta,\tilde{\eta})(f(u,\tilde{\eta}) - f(u,\eta)). \tag{1.1}$$

In this paper, we are particularly interested in the case when the differential flow generated by $\psi(\cdot,\eta)$ is contracting for each fixed η. Although this assumption seems trivial at first glance, the underlying dynamics is actually rich, because the attractor of $\psi(\cdot,\eta)$ is different for each η. Hence the ODE part $\mathbf{U}_t$ is dragged towards different points along the jumps of η_t, while influencing the transition rate of η_t at the same time. See Section 5.2 of [3] for a simple example. In principal, piecewise contraction holds for systems that are otherwise contracting if there is no stochastic turbulence. This includes a wide range of models in engineering, economics and natural science. On the other hand, it is relatively easy to verify as it depends solely on ψ.

Heuristically speaking, this paper proves the geometric ergodicity of the joint process $(\mathbf{U}_t,\eta_t)$ in a Wasserstein distance as long as the following three conditions are verified:

- There is a Lyapunov function for the system, and the transition rate with its Frechét derivative with respect to $\mathbf{U}$ are controlled by this Lyapunov function;
- The process is irreducible: there is a common state that is accessible from other states;
- Piecewise contracting: with the stochastic part η being fixed, the vector field $\psi(\cdot,\eta)$ is contracting.

The proof of the main theorem utilizes the differential flow structure of PDMP and invokes a perturbation analysis over the probability measure. The bounds generated by these analyses can be applied to the powerful asymptotic coupling framework developed in [16, 18], which finishes the proof.

2 Stochastic lattice models for the tropics

In this section we describe two stochastic lattice models, the simplest tropical climate model developed in [15, 32, 29] and the stochastic skeleton model for the MJO developed and applied in [33, 34, 40, 39]. They both consist of an ODE system $\mathbf{U}_t$, which describes the dry dynamics based on continuum thermal-dynamical PDEs, and a stochastic jump process η_t, which describes the intermittent tropical variability. In order to be consistent with this paper's emphasis, we are presenting only a minimal introduction of both models. Interested readers are referred to the corresponding references for systematic derivations and discussions of these models.

2.1 The simplest tropical climate model

Deterministic model

In [15, 32, 29], the simplest tropical climate model is derived to capture the impact of tropical moisture variability. Here we discuss a simplified setup of flows above the equator, which follows a PDE:

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial \theta}{\partial x} - \bar{d}u, \\ \frac{\partial \theta}{\partial t} - \frac{\partial u}{\partial x} &= -d_\theta(\theta - \theta_{eq}) + d_{sh}(\theta_s - \theta) + P, \\ \frac{\partial q}{\partial t} + \bar{Q}\frac{\partial u}{\partial x} &= d_q(q_s - q) - P. \end{aligned} \tag{2.1}$$

Here the periodic non-dimensional variable x denotes the longitude. Scalar fields u, θ and q denote the zonal velocity, the potential temperature and the moisture of the flow. $\theta_s, \theta_{eq}, q_s$ are fixed periodic function of x, while the damping coefficients d_q, d_{sh}, d_q represent the radiative cooling, the sensible heat flux and the evaporation. The precipitation P is modeled throughout Betts and Miller's method [5, 6] using the convective available potential energy (CAPE) [38]:

$$P = \tau_c^{-1}(q - \alpha\theta - \hat{q})^+$$

As physical constraints derived in [15], here we require that $\alpha \geq 0, 1 > \bar{Q} > 0$. In order to turn (2.1) into a numerically implementable model, and to mimic the coarse graining procedure of the classical general circulation model (GCM), we will apply spatial discretization. Consider a change of variables based on the Riemann invariants of (2.1) following [15]:

$$K = u - \theta, \quad R = -u - \theta, \quad Z = \bar{Q}\theta + q$$

then (K, R, Q) follows the following PDE:

$$\begin{aligned} \frac{\partial K}{\partial t} + \frac{\partial K}{\partial x} &= -\frac{\bar{d} + d_\theta + d_{sh}}{2}K - \frac{d_\theta + d_{sh} - \bar{d}}{2}R - (d_\theta\theta_{eq} + d_{sh}\theta_s + P), \\ \frac{\partial R}{\partial t} - \frac{\partial R}{\partial x} &= -\frac{\bar{d} + d_\theta + d_{sh}}{2}R - \frac{d_\theta + d_{sh} - \bar{d}}{2}K - (d_\theta\theta_{eq} + d_{sh}\theta_s + P), \\ \frac{\partial Z}{\partial t} &= -d_q Z + \frac{d_\theta + d_{sh} - d_q}{2}\bar{Q}(K + R) + \bar{Q}(d_\theta\theta_{eq} + d_{sh}\theta_s) + d_q q_s - (1 - \bar{Q})P. \end{aligned}$$

Applying the first order upwind numerical scheme, cf. [19], we discretize the PDE above into an ODE system on the coarse grained periodic lattice

$I := \mathbb{Z}/N\mathbb{Z}$:

(2.2)

$$\frac{dK_i}{dt} + D_x^+ K_i = -\frac{\bar{d}+d_\theta+d_{sh}}{2}K_i - \frac{d_\theta+d_{sh}-\bar{d}}{2}R_i - (d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i}+P_i),$$

$$\frac{dR_i}{dt} - D_x^- R_i = -\frac{\bar{d}+d_\theta+d_{sh}}{2}R_i - \frac{d_\theta+d_{sh}-\bar{d}}{2}K_i - (d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i}+P_i),$$

$$\frac{dZ_i}{dt} = -d_qZ_i + \frac{d_\theta+d_{sh}-d_q}{2}\bar{Q}(K_i+R_i) + \bar{Q}(d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i}) + d_qq_{s,i} - (1-\bar{Q})P_i.$$

Here the finite difference operators are defined as:

$$D_x^+ f_i = N^{-1}(f_i - f_{i-1}), \quad D_x^- f_i = N^{-1}(f_{i+1} - f_i).$$

Such difference scheme will retain the L^2-stability of the system due to Lemma 5.1 below. As an approximation, the original prognostic variables u, θ, q at site $x = i/N$ should have the value

$$u_i = \frac{1}{2}(K_i - R_i), \quad \theta_i = -\frac{1}{2}(K_i + R_i), \quad q_i = Z_i + \frac{\bar{Q}}{2}(K_i + R_i).$$

Stochastic parameterization

Convective inhibition (CIN) is induced by the negative potential energy over the vertical motion and exists in the equilibrium state. Its highly fluctuating behavior at sub-grid scales can be best described by an interacting particle system coupled with the thermal dynamical PDE above [31, 28, 30]. By coarse-graining this interacting particle system [28, 25], each group of l refined scale CIN sites are represented by their sum $\eta_i(t)$, which is shown to be a birth-death process with absorption and desorption rates as:

$$c_a(\eta_i) = \frac{l-\eta_i}{\tau_I}, \quad c_d(\eta_i) = \frac{\eta_i}{\tau_I}\exp\left(-2U_0\frac{\eta_i-1}{l-1} + \bar{\gamma}q_i - h_0\right). \tag{2.3}$$

The ODE system (2.2) affects the transition rates through the moisture $q_i = Z_i + \frac{\bar{Q}}{2}(K_i + R_i)$, since moisture decreases the potential for CIN, and in return increases precipitation. Since the precipitation is inhibited at the CIN cites, the overall precipitation at location i is modeled as:

$$P_i = \frac{l-\eta_i}{\tau_c l}(q_i - \alpha\theta_i - \hat{q})^+ = \frac{l-\eta_i}{\tau_c l}\left(Z_i + \frac{\alpha+\bar{Q}}{2}(K_i+R_i) - \hat{q}\right)^+. \tag{2.4}$$

Jointly, the process consists of an ODE system $\mathbf{U} = (K_i, R_i, Z_i)_{i\in I} \in \mathbb{R}^{3I}$ and a jump process $\eta_t = (\eta_i(t))_{i\in I} \in \{0,\dots,l\}^I$. The formal generator $\mathscr{L}$, (1.1), is

$$\mathscr{L}f(\mathbf{u},\eta) = \nabla_u f(\mathbf{u},\eta)\cdot\psi(\mathbf{u},\eta) + \sum_{i\in I} c_a(\eta_i)(f(\mathbf{u},\eta+e_i) - f(\mathbf{u},\eta))$$
$$+ \sum_{i\in I} c_d(\eta_i)(f(\mathbf{u},\eta-e_i) - f(\mathbf{u},\eta)),$$

where ψ is the vector field generated by (2.2) and $\eta \pm e_i$ is adding/subtracting η with 1 on lattice point i.

Geometric ergodicity

One of the applications of the general result, Theorem 3.10 in the next Section, is showing that the joint process $(\mathbf{U}_t, \eta_t)$ is geometrically ergodic as long as the differential flow generated by (2.2) is contracting for each fix η:

Theorem 2.1. *Assume the physical coefficients in* (2.2) *satisfy the following relation:*

$$(1-\bar{Q})(\alpha+\bar{Q})d_q(d_\theta+d_{sh}) \geq (d_\theta+d_{sh}-d_q)^2\bar{Q}^2. \tag{2.5}$$

then the simplest tropical climate model given by (2.2), (2.4) *and* (2.3) *is geometrically ergodic under a suitable Wasserstein distance.*

The geometric ergodicity and Wasserstein distance will be defined in detail in Section 3 and Theorem 3.10. The mild assumption over the coefficients here holds for most models in [28, 15, 29] and especially when $d_\theta + d_{sh} = d_q$, since $0 < \bar{Q} < 1$ and $\alpha \geq 0$.

2.2 Stochastic skeleton model for the MJO

The deterministic skeleton model

The deterministic skeleton model is derived in [33] to capture the intermittent and wave train features of the Madden-Julian oscillation (MJO) in the tropics. The simplest way to describe it is through the following PDE of the equatorial Kelvin and Rossby wave, K and R, the nonnegative convective activity envelope strength A, and the first vertical baroclinic, meridional Hermite mode of the moisture Q [33, 40]:

$$\begin{aligned}
\partial_t K + \partial_x K &= (S^\theta - \bar{H}A)/2 - \bar{d}K,\\
\partial_t R - \partial_x R/3 &= (S^\theta - \bar{H}A)/3 - \bar{d}R,\\
\partial_t Q + \bar{Q}(\partial_x K - \partial_x R/3) &= (\bar{H}A - S^\theta)(\bar{Q}/6 - 1) - \bar{d}Q,\\
\partial_t A &= \Gamma Q A.
\end{aligned} \tag{2.6}$$

Here all the variables are periodic functions with respect to the longitude variable x. S^θ is a nonnegative prescribed periodic function of x representing the external source of heating and moistening, and $\bar{d}$ is an arbitrarily small dissipation. The corresponding zonal, meridian, vertical velocity (u, v, w), potential temperature θ, moisture q, and pressure p can be recovered through:

$$\begin{gathered}
u = \cos z[(K-R)\phi_0 + R\phi_2/\sqrt{2}], \quad v = 4\cos z(\partial_x R + S^\theta - \bar{H}A)\phi_1/3\sqrt{2},\\
\theta = -\sin z[(K+R)\phi_0 + R\phi_2/\sqrt{2}], \quad q = Q\phi_0 \sin z,\\
p = \cos z[(K+R)\phi_0 + R\phi_2/\sqrt{2}], \quad w = (\partial_x u + \partial_y v)\sin z,
\end{gathered}$$

where the ϕ_i's are the L^2-basis based on Hermite functions:

$$\phi_0 = \sqrt{2}(4\pi)^{-1/4}\exp(-y^2/2),$$
$$\phi_1 = 2y(4\pi)^{-1/4}\exp(-y^2/2),$$
$$\phi_2 = (2y^2-1)(4\pi)^{-1/4}\exp(-y^2/2).$$

By applying the Riemann invariant, we will replace Q with

$$Z = Q - \bar{Q}(K+R),$$

and thus the penultimate equation of (2.6) is replaced by:

$$\partial_t Z = (S^\theta - \bar{H}A)(1-\bar{Q}) - \bar{d}Z.$$

Discretization and stochastic parametrization

Using the first order upwind scheme again, we discretize the Kelvin and Rossby formulation (2.6) into the following ODEs with i belongs to the periodic lattice $I := \mathbb{Z}/N\mathbb{Z}$:

$$\begin{aligned} &\frac{dK_i}{dt} + D_x^+ K_i = (S_i^\theta - \bar{H}A_i)/2 - \bar{d}K_i, \\ &\frac{dR_i}{dt} - D_x^- R_i/3 = (S_i^\theta - \bar{H}A_i)/3 - \bar{d}R_i, \\ &\frac{dZ_i}{dt} = (S_i^\theta - \bar{H}A_i)(1-\bar{Q}) - \bar{d}Z_i. \end{aligned} \tag{2.7}$$

In [40], the corresponding convective envelope is modeled by $A_i = \Delta A \eta_i$ with η_i being a birth-death process with absorption-desorption rates:

$$c_a(\eta_i, Q_i) = \begin{cases} \Gamma|Q_i|\eta_i + 1_{\eta_i=0} & Q_i \geq 0 \\ 1_{\eta_i=0} & Q_i < 0 \end{cases}, \quad c_d(\eta_i, Q_i) = \begin{cases} 0 & Q_i \geq 0 \\ \Gamma|Q_i|\eta_i & Q_i < 0 \end{cases}, \tag{2.8}$$

here $Q_i = Z_i + \bar{Q}(K_i + R_i)$. Notice that this choice is an approximation of the deterministic relation $\partial_t A_t = \Gamma Q A$ in (2.6) as its formal generator is:

$$\mathscr{L}A_i = \Delta A \mathscr{L}\eta_i = \Gamma Q_i A_i + \Delta A 1_{A_i=0}.$$

In summary, the joint process consists of an ODE system $\mathbf{U} = (K_i, R_i, Z_i)_{i\in I} \in \mathbb{R}^{3N}$ and a jump process $\eta_t = (\eta_i(t))_{i\in I} \in \mathbb{Z}_+^N$. The formal generator $\mathscr{L}$, (1.1), is given by

$$\begin{aligned} \mathscr{L}f(\mathbf{u},\eta) = \nabla_u f(\mathbf{u},\eta)\cdot\psi(\mathbf{u},\eta) &+ \sum_{i=1}^N c_a(\eta_i)(f(\mathbf{u},\eta+e_i) - f(\mathbf{u},\eta)) \\ &+ \sum_{i=1}^N c_d(\eta_i)(f(\mathbf{u},\eta-e_i) - f(\mathbf{u},\eta)), \end{aligned}$$

where ψ is the vector field generated by (2.7) and $\eta \pm e_i$ is adding/subtracting η with 1 on lattice point i.

Geometric ergodicity

The jump process η_t of the skeleton model is very complicated because:

1) it takes place in an infinite space $\mathbb{Z}_+^I$;
2) the transition rates (2.8) are unbounded;
3) the transition rates (2.8) are degenerate, i.e. not strictly positive, especially at $Q_i = 0$.

In particular, there is a possible degeneracy here: suppose $S_i^\theta = \bar{H}\Delta A\eta_i$ for a group of integers η_i, then $(\vec{0}, \eta)$ is a fixed point for jump process, since the differential flow is stopped, i.e. $\psi(\vec{0}, \eta) = \vec{0}$, and all the transition rates are 0 because $Q_i = 0$. This will render the system either non-ergodic or the equilibrium state is not interesting. On the other hand, this degeneracy can be easily ruled out for most real applications due to the following lemma:

Lemma 2.2. *If $\sum S_i^\theta$ is not an integer multiple of $\bar{H}\Delta A$, there is no joint state $(\mathbf{u}, \eta) \in \mathbb{R}^{3I} \times \mathbb{Z}_+^I$ such that $\psi(\mathbf{u}, \eta) = 0$ and $Q_i = 0$ for all i.*

Proof. Assume the opposite, using that $Q_i = 0$ and the formulation of (2.7), we find

$$0 = \frac{d}{dt}\sum_i (\bar{Q}K_i + \bar{Q}R_i + Z_i) = (1 - \bar{Q}/6)\sum_i (S_i^\theta - \bar{H}A_i)$$

Since $\bar{Q} < 1$, the coefficient above before $\sum_i (S_i^\theta - \bar{H}A_i)$ is non zero, hence $\sum_i S_i^\theta = \bar{H}\Delta A \sum_i \eta_i$, which implies $\sum S_i^\theta$ is an integer multiple of $\bar{H}\Delta A$, contradicting our assumption. $\square$

On the other hand, as long as we rule out the possibility of a fixed point for the joint process, the skeleton model of MJO is geometric ergodic under a Wasserstein distance:

Theorem 2.3. *Assume that $\sum S_i^\theta$ is not an integer multiple of $\bar{H}\Delta A$, then the skeleton model of MJO, i.e. the process $(\mathbf{U}_t, \eta_t)$ with evolution given by* (2.7) *and* (2.8)*, is geometric ergodic under a suitable Wasserstein distance.*

Again, the geometric ergodicity and the Wasserstein distance will be defined in detail in Sections 3 and 6.

3 Geometric ergodicity for piecewise contracting systems

The stochastic lattice models set up in Section 2 have the same mathematical structure: there is an ODE system $\mathbf{U}_t$ and a Markov jump process η_t. Such processes are known as piecewise deterministic Markov processes (PDMP), or more generally Markov processes with random switching. The classical ergodicity results usually require the transition rates of η_t to be constants independent of $\mathbf{U}_t$ [35, 1, 2]. In recent years, there is a growing interest in extending these results to non-constant rates, either through hypoelliptic conditions [3], or through Wasserstein contraction [4, 7]. However, these results all require the transition rates of η_t to be bounded and globally Lipschitz, and produces a non-degenerate jump chain. Unfortunately, the stochastic lattice models introduced in Section 2 do not satisfy these conditions, especially the skeleton model of MJO based on the discussion of Section 2.2. The main objective here is to develop a theoretical framework using Lyapunov functions so unbounded and degenerate transition rates are allowed.

3.1 Definitions and notations

A PDMP can be defined as follows when there is an underlying differential flow structure:

Definition 3.1. *Let X_t be a continuous process in a Hilbert space $\mathscr{H}$ with norm $\|\cdot\|$, and Y_t be a cadlag (continuous from right with limits from left) process taking value in a countable set F. We call the joint process $Z_t = (X_t, Y_t) \in E := \mathscr{H} \times F$ a piecewise deterministic Markov process (PDMP) if the following holds:*

(1) Given the realization of $Y_{s\le t}$, X_t follows a non-explosive differential flow generated by a locally Lipschitz vector field ψ :

$$X_t = X_0 + \int_0^t \psi(X_s, Y_s)ds =: \Psi(X_0, Y_{s\le t}, t) \quad \forall t \ge 0.$$

When the process Y_s takes constant value y, the corresponding trajectory of X_t will be denoted as $\Psi_s^y x$, which is the solution to the equation: $\Psi_t^y x = x + \int_0^t \psi(\Psi_s^y x, s)ds$.

(2) Given the value of X_t, Y_t is a continuous time Markov chain with jump intensity from current state to $y' \in F$ being $\lambda_t(X_t, Y_{t-}, y')$. Here $Y_{t-} := \lim_{s\nearrow t} Y_s$. We also write the sum of jump rates as: $\bar{\lambda}(x,y) := \sum_{y'\in F} \lambda(x,y,y')$. In other words,

$$\mathbb{P}(Y_s \text{ does not jump in } [0,t] | X_0 = x, Y_0 = y) = \exp\left(-\int_0^t \bar{\lambda}(\Psi_s^y x, y)ds\right),$$

$$\mathbb{P}(Y_t = y' | Y_t \text{ jumps at time } t, X_t, Y_{t-}) = \frac{\lambda(X_t, Y_{t-}, y')}{\bar{\lambda}(X_t, Y_{t-})}.$$

The formal construction of such process and the verification that they are Markovian can be found in [10, 20], while a shorter self explanatory version can be found on page 3 of [7].

Remark 3.2. *PDMP can also be defined for more general models, where Ψ can explode in finite time or X_t takes values in a family of spaces indexed by F. Interested readers are directed to* [10, 20] *for these extensions. We use this more practical definition of PDMP in order to avoid unnecessary abstractions.*

For the notation in this section and the next, we use symbol Z to denote the joint process (X_t, Y_t), while X is its differential flow part and Y is its random jump process part. We save the symbols $\mathbf{U}_t$ and η_t for applications in Section 5 and 6. We will write the transition rate in two different fashions in order to emphasize different variables:

$$\lambda(z,k) = \lambda(x,y,k), \quad \bar{\lambda}(z) = \bar{\lambda}(x,y)$$

We also use $\mathbb{P}^z$ to denote the law of $Z_{s\geq 0}$ given that $Z_0 = z = (x,y)$ and $\mathbb{E}^z$ to denote the corresponding expectation. $\mathbb{P}^\mu$ denotes the law of $Z_{s\geq 0}$ with $Z_0 \sim \mu$. Thus the generated transition kernel is denoted as P_t, that is $P_t f(z) := \mathbb{E}^z f(Z_t)$.

We use $\tau_1, \dots, \tau_k$ to denote the jump times of Y_t, and N_t denotes the total number of jumps up to time t. The differential flow $\Psi(x, Y_{s\leq t}, t)$ can as well be written as:

$$X_t = \Psi(x, Y_{s\leq t}, t) = \Psi^{Y_{\tau_{N_t}}}_{t-\tau_{N_t}} \Psi^{Y_{\tau_{N_t-1}}}_{\tau_{N_t}-\tau_{N_t-1}} \cdots \Psi^{Y_0}_{\tau_1} x.$$

Finally, let us define the formal generator $\mathscr{L}$ for functions on $\mathscr{H} \times F$ that are C^1 in x:

$$\mathscr{L} f(x,y) = \psi(x,y)\partial_x f(x,y) + \sum_{y'\in F} \lambda(x,y,y')(f(x,y') - f(x,y)). \tag{3.1}$$

Remark 3.3. *Evidently, $\mathscr{L}$ is the candidate for the infinitesimal generator. When the total jump intensity is uniformly bounded, it is known that $\mathscr{L}$ is the generator for the process. See Lemma 2.1 of* [3] *or Remark 26.16 of* [10]. *This gives the useful Dynkin's formula for all bounded $f(x,y)$ that are C^1 in x:*

$$\mathbb{E}^z f(Z_t) = f(z) + \mathbb{E}^z \int_0^t \mathscr{L} f(Z_s) ds. \tag{3.2}$$

Rigorously speaking, $\mathscr{L}$ will be the extended operator with its domain given by Theorem 26.14 of [10]. *Yet, we are not introducing these terminologies here as they are quite abstract. On the other hand, our applications of $\mathscr{L}$ are restricted to a few simple functions, where the properties of the generator follow easily over a localization sequence, see Lemma 3.7 below for example.*

3.2 Stability and Regularity

The Definition 3.1 of PDMP is rather general so there is no guarantee for any notions of stability. In fact, it is even possible to have infinite jumps in Y_t in finite time as the state space E is non-compact and the total jump intensity is unbounded. One standard tool to stabilize a system in non-compact space is the Lyapunov function [37, 36, 16, 18, 7], which will play a crucial role in our theoretical framework:

Assumption 3.4 (Lyapunov Function)**.** *There exists a function $V(x,y)$: $E \mapsto \mathbb{R}_+$, which is C^1 with respect to x, with sub-level sets being compact and the following hold with some $\gamma, k_v > 0$:*

$$\|x\|^2 \le V(z), \quad \mathscr{L}V(z) \le -\gamma V(z) + k_v.$$

Notice that if V satisfies such conditions, then so does $V+1$. Hence without lost of generality, we assume $V \ge 1$.

Remark 3.5. *Another more general way to define Lyapunov function is requiring that $\mathbb{E}^z V(Z_t) \le e^{-\gamma t} V(z) + K_v$, as in* [7]. *The generator form of Lyapunov function we propose here and also in* [36] *is slightly stronger due to Lemma 3.7 in below. We choose this form because at this stage it is unclear whether Y_t will explode in finite time, so it is problematic to use $\mathbb{E}^z V(Z_t)$. However, $\mathscr{L}$ can always be formally defined as* (3.1).

This Lyapunov function V will bound and regularize the transition rates in the following manner:

Assumption 3.6 (Regularity of Transition rates)**.** *The transition rates have the following properties for a proper constant $M_\lambda > 0$ with any $z \in E, y' \in F$:*

(1) $\lambda(z,y') \neq 0$ *for at most M_λ different $y' \in F$;*
(2) $\bar{\lambda}(z), \|D_x \lambda(x,y,y')\|, \|D_x \overline{\lambda}(x,y)\| \le M_\lambda V(x,y)$. *Here and after, D_x denotes the Frechét derivative with respect to the x.*

Assumption 3.4 and 3.6 in combination insure no explosion, and $\mathscr{L}$ works as a generator on V, as shown by the following lemma:

Lemma 3.7. *With Assumptions 3.4 and 3.6, then $\mathbb{P}^z$-a.s. there is no explosion and for some $K_t < \infty$*

$$\mathbb{E}^z N_t = \mathbb{E}^z \int_0^t \bar{\lambda}(Z_{s-}) ds \le K_t V(z), \quad \mathbb{E}^z V(Z_t) \le e^{-\gamma t} V(z) + k_v/\gamma.$$

Proof. Fix any $n \in \mathbb{N}$, consider stopping times $T_n := \inf\{t : V(Z_t) > n\}$. The corresponding process with its associated total jumps can be jointly written as:

$$(\widetilde{Z}^n_t, \widetilde{N}^n_t) := (X_{t\wedge T_n}, (Y_{t\wedge T_n}, N_{t\wedge T_n})).$$

It can be easily verified as a PDMP that takes place in $\mathscr{H} \times (F \times \mathbb{N})$. Its differential flow is generated by the vector field $\psi(z) 1_{V(z)<n}$, and the jump rate from state (y,m) to $(y', m+1)$ with the X part being x is $\lambda(x,y,y') 1_{V(z)<n}$. Such rates by Assumption 3.6 are uniformly bounded, so following Remark 3.3 and the reference within, we know its generator is:

$$\begin{aligned}\mathscr{L}^n f(x,y,m) =& 1_{V(z)<n} \psi(x,y) \partial_x f(x,y,m) \\ &+ 1_{V(z)<n} \sum_{y' \in F} \lambda(x,y,y')[f(x,y',m+1) - f(x,y,m)].\end{aligned}$$

Apply the Dynkin's formula (3.2) to $\widetilde{N}_t^n = \pi(\widetilde{Z}_t^n, \widetilde{N}_t^n)$, where $\pi(z,m) = m$:

$$\mathbb{E}^z N_{t\wedge T_n} = \mathbb{E}^z \int_0^{t\wedge T_n} \bar{\lambda}(Z_s)ds \leq M_\lambda \mathbb{E}^z \int_0^{t\wedge T_n} V(Z_s)ds. \tag{3.3}$$

On the other hand, $\exp(\gamma t \wedge T_n)V(\widetilde{Z}_t^n)$ is bounded for bounded t, so the Dynkin's formula (3.2) gives:

$$\mathbb{E}^z \exp(\gamma(t\wedge T_n))V(\widetilde{Z}^n_{t\wedge T_n}) = V(z) + \mathbb{E}^z \int_0^{t\wedge T_n} \mathscr{L}^n \exp(\gamma s)V(\widetilde{Z}_s^n)ds.$$

By the chain rule and Assumption 3.4, when $V(z) < n$,

$$\mathscr{L}^n \exp(\gamma t)V(z) = \exp(\gamma t)(\gamma V(z) + \mathscr{L}V(z)) \leq \exp(\gamma t)k_v.$$

Since before T_n, $V(z) < n$, so the following holds

$$\mathbb{E}^z \exp(\gamma(t\wedge T_n))V(\widetilde{Z}^n_{t\wedge T_n}) \leq V(z) + \exp(\gamma t)k_v/\gamma. \tag{3.4}$$

Notice that $\mathbb{E}^z V(Z_{t\wedge T_n}) \leq \mathbb{E}^z \exp(\gamma(t\wedge T_n))V(\widetilde{Z}^n_{t\wedge T_n})$, so by the Markov inequality:

$$\mathbb{P}(T_n < t) \leq n^{-1}\mathbb{E}^z V(Z_{t\wedge T_n}) \to 0.$$

By Borel-Cantelli's lemma, there is a proper subsequence, $\{n_k\}$, such that

$$T_{n_k} \to \infty, \quad \exp(\gamma(t\wedge T_{n_k}))V(\widetilde{Z}^{n_k}_{t\wedge n_k}) \nearrow \exp(\gamma t)V(Z_t) \quad \mathbb{P}^z\text{-a.s.}$$

Apply Fatou's lemma over this subsequence to (3.4), we have

$$\mathbb{E}^z \exp(\gamma t)V(Z_t) \leq V(z) + \exp(\gamma t)k_v/\gamma.$$

This generates the second inequality in the statements of this lemma. Apply Fatou's lemma over the same subsequence to (3.3), we have

$$\mathbb{E}^z N_t = \mathbb{E}^z \int_0^t \bar{\lambda}(Z_s)ds \leq M_\lambda \mathbb{E}^z \int_0^t V(Z_s)ds \leq M_\lambda [V(z)/\gamma + tk_v/\gamma].$$

This implies the first inequality in the statements of this lemma and $N_t < \infty$ a.s. $\square$

.

3.3 Piecewise contraction and accessibility

There are two more conditions we need to show a PDMP is geometrically ergodic.

Assumption 3.8 (Piecewise Contraction)**.** *The following holds for some proper constants $C_\gamma, \gamma > 0$:*

$$\|D_x\Psi(x, y_{s\leq t}, t)\| \leq C_\gamma \exp(-\gamma t).$$

for all F-valued process $y_{s\leq t}$ and $x \in \mathscr{H}$. As a special case, the flow with $y_{s\leq t}$ being of constant value y, $\Psi_t^y x$, will converge to a single point for each fixed y (maybe infinity), which will be called the attractor for state y in the following.

The final assumption over the model is the irreducibility of the model, phrased by the accessibility:

Assumption 3.9 (Accessibility from compact sets)**.** *For any compact subset C of E, there exists a $y_c \in F$ such that its corresponding attractor x_c satiefies $V(x_c, y_c) < \infty$ and for any $z \in C$, there exists a t such that $\mathbb{P}^z(Y_t = y_c) > 0$. Here the topology of $\mathscr{H}$ is generated by its norm $\|\cdot\|$, the topology of F is discrete and the topology of E is the product of the two.*

The classical notion of ergodicity is usually illustrated in the total variation norm. Results like $\|\mathbb{P}^z_t - \mathbb{P}^{z'}_t\|_{TV} \xrightarrow{t\to\infty} 0$ are well studied and understood in the finite dimensional Markov chain or stochastic Markov process setting by [37, 36]. Yet, for PDMP, convergence in total variation maybe too stringent as the total variation distance discriminates deterministic systems rather harshly. For example, consider a deterministic process in $\mathbb{R}$, $dX_t = -\gamma X_t dt$. The invariant measure is obviously δ_0, a point mass at the origin. Yet, starting from any nonzero point, the distribution of X_t is a point mass at $e^{-\gamma t}X_0$, which has total variation distance 2 from δ_0. One way to guarantee convergence in total variation, is to assume Hörmander type of conditions over the vector fields [2, 3]. However checking Hörmander condition is tedious for ODE systems on lattice structure when there are neighboring interaction, for example the stochastic lattice models introduced in Section 2.

A more suitable distance between measures for PDMP, is the Wasserstein-1 distance, which is also used in previous works for PDMP [1, 4, 7]. For any distance d on E, the Wasserstein distance with respect to d between two measures μ, ν on E is defined as:

$$d(\mu,\nu) := \inf_{\Gamma\in\mathscr{C}(\mu,\nu)} \int d(x,x')\Gamma(dx,dx')$$

Here $\mathscr{C}(\mu,\nu)$ is the set of all couplings between μ and ν; in other words the marginal distributions of any $\Gamma \in \mathscr{C}(\mu,\nu)$ are μ and ν respectively.

Theorem 3.10. *Suppose Assumptions 3.4, 3.6, 3.8 and 3.9 hold, then the PDMP $Z_t = (X_t, Y_t)$ has a unique invariant measure π, moreover the distribution of Z_t converges to π geometrically fast in the Wasserstein distance with respect to the natural distance*

$$m(z,z') = \sqrt{\|x-x'\|^2 + 1_{y\neq y'}}.$$

In particular, there are some proper constants $\beta, C > 0$ such that the following holds for any $z, z' \in E$:

$$m(\mathbb{P}^z_t, \mathbb{P}^{z'}_t) \le C\exp(-\beta t)\sqrt{1+V(z)+V(z')}.$$

4 Proofs

The proof of Theorem 3.10 uses the asymptotic coupling mechanism introduced by [18, 17]. Theorem 4.8 of [18] gives the major application of this mechanism.

Theorem 4.1 (HMS11). *Let P_t be a Markov semigroup over a Polish space E admitting a continuous Lyapunov function $V: P_tV(z) \leq C_V e^{-\gamma t}V(z) + K_V$. Suppose furthermore that there exists $t > 0$ and a distance-like function $d: E \times E \mapsto [0,1]$ which is contracting for P_t and such that the level set $\{z \in E : V(z) \leq 4K_V\}$ is d-small for P_t. Then P_t can have at most one invariant probability measure π. Furthermore, defining*

$$\tilde{d}(z,z') = \sqrt{d(z,z')(1+V(z)+V(z'))},$$

there exists $t > 0$ such that $\tilde{d}(\mathbb{P}_t^\mu, \mathbb{P}_t^\nu) \leq \frac{1}{2}\tilde{d}(\mu,\nu)$ for any probability measures μ,ν on E.

In [18] a distance d is $\frac{1}{2}$-contracting for Markov chain $(Z_{nt})_{n=1,\dots}$ if $d \leq 1$ and

$$d(\mathbb{P}_t^z, \mathbb{P}_t^{z'}) \leq \frac{1}{2}d(z,z'), \quad \forall d(z,z') < 1.$$

The distance between measures are understood as the Wasserstein distance with respect to d introduced before Theorem 3.10. And a set A is d-small for the chain $(Z_{nt})_{n=1,\dots}$ if there exists an $\varepsilon > 0$ such that

$$d(\mathbb{P}_t^z, \mathbb{P}_t^{z'}) \leq 1-\varepsilon, \quad \forall z,z' \in A.$$

We will adopt these concepts in the following. Evidently, the key step in applying Theorem 4.1 is constructing a contracting distance d. Proposition 5.5 of [18] sets up a versatile framework for this purpose, and we will use a variant of it to set up a contracting distance. The validation of this framework relies on an analysis in the perturbation of the underlying probability. In the following, we will first study the probability density of process Z_t in Section 4.1, then analyze the perturbation of this density and hence construct the contracting distance in Section 4.2, while Sections 4.3 and 4.4 use Theorem 4.1 to prove Theorem 3.10.

4.1 Probabilistic density for admissible jumps

As the differential flow of a PDMP is deterministic, the path of $Z_t = (X_t, Y_t)$ solely depends on the realization of the number of jumps N_t, the jump times τ_i and destinations of the jumps Y_{τ_i}. In fact, the whole path can be written down explicitly as a function of the sequence $\mathbf{t} = (t_1, \dots, t_n), \mathbf{y} = (y_{t_1}, \dots, y_{t_n})$:

$$(4.1)\quad y_s = \begin{cases} y_{t_k}, & s \in [t_k, t_{k+1}) \\ y_{t_n}, & s \geq t_n \end{cases} \quad ; \quad x_s = \Psi(x_0, \mathbf{t}, \mathbf{y}, s) := \Psi(x_0, y_{r\leq s}, s).$$

According to the construction of PDMP in Definition 3.1, using a standard Kolmogorov procedure one can derive the following probability density of $(n,\mathbf{t},\mathbf{y}) \in \mathbb{N}\times\mathbb{R}_+^n\times F^n$. This is carried out in detail by Theorem 7.3.1 and formula 3.10 of [20].

(4.2)

$$\begin{aligned}&\mathbb{P}(N_t=n, Y_{\tau_1}=y_{t_1},\dots,Y_{\tau_n}=y_{t_n},\tau_1\in dt_1,\dots,\tau_n\in dt_n)\\ &\quad= 1_{t_1<\dots<t_n}\exp\left(-\int_0^t\bar\lambda(z_s)ds\right)\prod_{i=1}^n\left(\lambda(x_{t_i},y_{t_{i-1}},y_{t_i})dt_i\right) =: p^{z,t}_{n,\mathbf{t},\mathbf{y}}d\mathbf{t}.\end{aligned}$$

The process $z_s=(x_s,y_s)$ in the formula above is defined through (4.1) as a function of $(n,\mathbf{t},\mathbf{y})$. The underlying measure of the density $p^{z,t}_{n,\mathbf{t},\mathbf{y}}$ is the Lebesgue measure on $\mathbb{R}^n$: $d\mathbf{t}=dt_1\,dt_2\dots dt_n$. The intersection with event $\{N_t=n\}$ here is necessary for the definition of density, else the number of τ_i will be unclear and hence also the dimension of underlying Lebesgue measure. On the other hand, when using this formula to compute expectation, one must remember to enumerate among all possible values of n. Lemma 3.7 guarantees that there is no explosion, so the enumeration only need to go through $n\in\mathbb{N}$. In other words, for any measurable function f, its expectation is given by:

$$\mathbb{E}^z f(Z_{s\le t}) = \sum_{n=0}^{\infty}\sum_{\mathbf{y}\in F^n}\int_{[0,t]^n} d\mathbf{t} f(z_{s\le t})p^{z,t}_{n,\mathbf{t},\mathbf{y}}. \tag{4.3}$$

4.2 Contracting distance

For the construction of a contracting distance, we have the following lemma. It is a variant of Proposition 5.5 in [18]. Here we use a Lyapunov function instead of a super Lyapunov function and we need to deal with two processes X_t and Y_t instead of one. Yet, the nature of the proof remains quite the same.

Lemma 4.2. *Suppose that a transition kernel Q on space $E=\mathscr{H}\times F$ satisfies $QV(z)\le DV(z)$ with a constant D, while the following holds for any $\varphi: E\mapsto\mathbb{R}$ that is C^1 in $\mathscr{H}$:*

$$\|D_xQ\varphi(z)\| \le \left(\frac{1}{4D}[Q\|\partial_x\varphi\|(z)] + CV(z)\|\varphi\|_\infty\right). \tag{4.4}$$

Here, $\|f\|_\infty=\sup_z|f(z)|$ where $|\cdot|$ denotes a proper norm for f. Then there exists a $\delta>1$ such that if we define

$$d(z,z') = 1_{y\ne y'} + 1_{y=y'}\wedge\left(\delta^{-1}\inf_{r:x\to x'}\int_0^1 V(r(s),y))\|\dot r(s)\|ds\right), \tag{4.5}$$

d is a $\frac12$-contracting metric for the chain Z_n generated by Q. Here the infimum is taken over all C^1 path r such that $r(0)=x, r(1)=x'$.

Proof. Denote the law of z after transition Q as Q^z. By the definition of contracting metric, we need to show $d(Q^z,Q^{z'}) \le \frac{1}{2}d(z,z')$ when $d(z,z') < 1$, which implies that $y = y'$. Since the spaces here are Polish, by the Kantorovich-Rubinstein Theorem [12, Theorem 11.8.2],

$$d(Q^z,Q^{z'}) = \sup\left\{ \int \varphi dQ^z - \varphi dQ^{z'} \middle| \|\varphi\|_{Lip(d)} \le 1 \right\}.$$

where the requirement $\|\varphi\|_{Lip(d)} \le 1$ is equivalent to: for any $z,z' \in E$, $\varphi(z) - \varphi(z') \le d(z,z')$. Hence, to prove this lemma we only need to show that for any φ, $\|\varphi\|_{Lip(d)} \le 1$,

$$Q^{x,y}\varphi - Q^{x',y}\varphi \le \frac{1}{2\delta}\int_0^1 V(r(s),y)\|\dot r(s)\|ds. \tag{4.6}$$

However, if $\|\varphi\|_{Lip(d)} \le 1$, the maximum variation of φ is less than 1, hence we can replace φ by $\varphi - c$ such that $\|\varphi\|_\infty \le \frac{1}{2}$, yet $\int \varphi dQ^{x,y} - \varphi dQ^{x',y}$ remains invariant. So without loss of generality, we assume $\|\varphi\|_\infty \le \frac{1}{2}$.

Moreover, since d is equivalent to $\|\cdot\|$ in the x part, φ is continuous in x. If φ is not C^1 in x, we can find a sequence of C^1 functions φ_n such that $\varphi_n \to \varphi$ point wise. If we can show the upper bound holds for each φ_n, then since $Q^z\varphi_n \to Q^z\varphi$, with dominated convergence theorem we can show the upper bound holds for φ as well. So without loss of generality, we assume φ is C^1 in x.

Also note that ψ is d-Lipschitz implies that $\|\partial_x\varphi(z)\| \le \delta^{-1}V(z)$ since for any $v \in H, \|v\| = 1$,

$$\langle \partial_x\varphi, v\rangle \le \left| \lim_{\varepsilon\to 0} \frac{\varphi(x+\varepsilon v,y) - \varphi(x,y)}{\varepsilon} \right| \le \delta^{-1}V(z).$$

Therefore, it suffices for us to show the inequality (4.6) for φ that is

$$\|\varphi\|_\infty \le \frac{1}{2}, \quad \|\partial_x\varphi\| \le \delta^{-1}V(z).$$

Plug them into the assumption (4.4), we obtain,

$$\|D_xQ\varphi\| \le \left(\frac{1}{4D\delta}[QV(z)] + \frac{C}{2}V(z)\right) \le \left(\frac{1}{4\delta} + \frac{C}{2}\right)V(z).$$

Pick $\delta \le \frac{1}{2C}$, one has $\|D_xQ\varphi\| \le (2\delta)^{-1}V(z)$, which concludes our result because for any path r connecting x,x':

$$\begin{aligned}|Q\varphi(x,y) - Q\varphi(x',y)| &\le \int_0^1 \|D_xQ\varphi(r(s),y)\|\|\dot r(s)\|ds\\ &\le \frac{1}{2\delta}\int_0^1 V(r(s),y)\|\dot r(s)\|ds.\end{aligned}$$

Minimize over all paths generates our claim. $\square$

The next lemma verifies the condition of Lemma 4.2 under our setting.

Lemma 4.3. *With Assumptions 3.4, 3.6 and 3.8, there exists a T such that* (4.4) *holds for $Q = P_T$. Therefore d defined by* (4.5) *is $\frac{1}{2}$-contracting for P_T.*

Proof. First of all, according to Lemma 3.7 and the convention that $V \geq 1$,

$$P_t V(z) \leq V(z) + k_v/\gamma \leq (1 + k_v/\gamma) V(z), \quad \forall t \geq 0.$$

Hence $D = 1 + k_v/\gamma$ is good for the first condition in Lemma 4.2. Next, we verify (4.4) by taking Frechét derivative over the expectation formula (4.3). Using the chain rule, we have:

$$\begin{aligned}
\|D_x P_t \varphi(z)\| &= \left\| D_x \sum_{n,\mathbf{y}} \int d\mathbf{t} \varphi(\Psi(x,\mathbf{t},\mathbf{y},t), y_t) p^{z,t}_{n,\mathbf{t},\mathbf{y}} \right\| \\
&\leq \left\| \sum_{n,\mathbf{y}} \int d\mathbf{t} D_x[\varphi(\Psi(x,\mathbf{t},\mathbf{y},t), y_t)] p^{z,t}_{n,\mathbf{t},\mathbf{y}} \right\| \\
&\quad + \left\| \sum_{n,\mathbf{y}} \int d\mathbf{t} \varphi(\Psi(x,\mathbf{t},\mathbf{y},t), y_t) D_x(p^{z,t}_{n,\mathbf{t},\mathbf{y}}) \right\| \\
&\leq \left\| \sum_{n,\mathbf{y}} \int d\mathbf{t} \|\partial_x \varphi(z_t)\| \|D_x \Psi(x,\mathbf{t},\mathbf{y},t)\| p^{z,t}_{n,\mathbf{t},\mathbf{y}} \right\| \\
&\quad + \sum_{n,\mathbf{y}} \int d\mathbf{t} \|\varphi\|_\infty \|D_x(p^{z,t}_{n,\mathbf{t},\mathbf{y}})\| \\
&\leq C_\gamma e^{-\gamma t} \mathbb{E}^z \|\partial_x \varphi(Z_t)\| + \|\varphi\|_\infty \sum_{n,\mathbf{y}} \int d\mathbf{t} \|D_x p^{z,t}_{n,\mathbf{t},\mathbf{y}}\|. \qquad (4.7)
\end{aligned}$$

At the third step, we use $\|D_x \Psi(x,\mathbf{t},\mathbf{y},y)\| \leq C_\gamma e^{-\gamma t}$ in Assumption 3.8 to get (4.7). Pick a T such that $C_\gamma e^{-\gamma T} \leq 1/4D$. It remains to find a $C < \infty$ such that

$$\sum_{n,\mathbf{y}} \int d\mathbf{t} \|D_x p^{z,T}_{n,\mathbf{t},\mathbf{y}}\| \leq C V(z).$$

Apply the chain rule to the density $p^{z,T}_{n,\mathbf{t},\mathbf{y}}$ as in (4.2), we have

$$\begin{aligned}
\sum_{n,\mathbf{y}} \int d\mathbf{t} \|D_x p^{z,T}_{n,\mathbf{t},\mathbf{y}}\| &\leq \sum_{n,\mathbf{y}} \int d\mathbf{t} p^{z,T}_{n,\mathbf{t},\mathbf{y}} \int_0^T \|D_x \Psi(x,\mathbf{t},\mathbf{y},s)\| \|D_x \overline{\lambda}_s(x_s, y_s)\| ds \\
&\quad + \sum_{n,\mathbf{y}} \int d\mathbf{t} p^{z,T}_{n,\mathbf{t},\mathbf{y}} \sum_{k=1}^n \|D_x \Psi(x,\mathbf{t},\mathbf{y},t_k)\| \frac{\|D_x \lambda(z_{t_{t_k}-}, y_{t_k})\|}{\lambda(z_{t_k -}, y_{t_k})}. \qquad (4.8)
\end{aligned}$$

Notice here we write $z_{t_k -} = (x_{t_k}, y_{t_{k-1}})$. We will bound the two parts separately in the following. Denote A_z as the possible sites for the transition to happen from state $z = (x,y)$:

$$A_z := \{y' \in F : \lambda(z,y') > 0\}.$$

By Assumption 3.8 $\|D_x \Psi(x,\mathbf{t},\mathbf{y},s)\| \leq C_\gamma e^{-\gamma s}$, combine it with $\|D_x \bar{\lambda}(z)\| \leq M_\lambda V(z)$ from Assumption 3.6, and the first part of (4.8) can be bounded by:

$$\sum_{n,\mathbf{y}}\int d\mathbf{t}p^{z,T}_{n,\mathbf{t},\mathbf{y}}\int_0^T \|D_x\Psi(x,\mathbf{t},\mathbf{y},s)\|\|D_x\overline{\lambda}(x_s,y_s)\|ds$$
$$\leq C_\gamma M_\lambda\sum_{n,\mathbf{y}}\int d\mathbf{t}p^{z,T}_{n,\mathbf{t},\mathbf{y}}\int_0^T e^{-\gamma s}V(z_s)ds\leq C_\gamma M_\lambda\sum_{n,\mathbf{y}}\int d\mathbf{t}p^{z,T}_{n,\mathbf{t},\mathbf{y}}\int_0^T V(z_s)ds$$
$$=C_\gamma M_\lambda\mathbb{E}^z\int_0^T V(Z_s)ds\leq C_\gamma K_T T M_\lambda V(z),$$

using Lemma 3.7. For the second part of (4.8), first consider the following process which is a modification of Z_t at τ_k: let $\widetilde{Z}^k_t$ be the same process as Z_t before and after the k-th jump time τ_k, while at τ_k, the transition is uniform among the finite set $A_{Z_{\tau_k-}}\subset F$. Another way to define this process is letting $(\widetilde{Z}^k_t,N_t)$ to be a PDMP with formal generator:

$$\mathscr{L}f(z,n)=\psi(z)\partial_x f(z,n)+\sum_{y'\in A_z}\lambda(z,y')(f(x,y',n+1)-f(x,y,n)),\ n\neq k-1;$$

$$\mathscr{L}f(z,n)=\psi(z)\partial_x f(z,n)+\sum_{y'\in A_z}\frac{\bar\lambda(z)}{\#A_z}(f(x,y',n+1)-f(x,y,n)),\quad n=k-1.$$

Following (4.2), the probability density of this process up to time T can be written down as:

$$\tilde p^k_{n,\mathbf{t},\mathbf{y}}=1_{t_1<t_2<\ldots<t_n}\exp\left(-\int_0^T\bar\lambda(z_s)ds\right)\prod_{i\neq k}^n\lambda(z_{t_i-},y_{t_i})\left[\frac{\bar\lambda(z_{t_k-})}{\#A_{z_{t_k}-}}\right]^{1_{k\leq n}}.$$

Denote the law of $\widetilde{Z}^k_t$ as $\mathbb{P}^k$, notice that it coincides with $\mathbb{P}$ until τ_k. By Assumption 3.6, $\#A_z\leq M_\lambda$, so comparing with the formula (4.2), we have:

$$M_\lambda^2\frac{V(z_{t_{k-}})}{\bar\lambda(z_{t_{k-}})}\tilde p^k_{n,\mathbf{t},\mathbf{y}}\geq p^{z,t}_{n,\mathbf{t},\mathbf{y}}\frac{\|D_x\lambda(z_{t_k-},y_{t_k})\|}{\lambda(z_{t_k-},y_{t_k})}.$$

Hence the second part of (4.8) can be bounded by:

$$\sum_{n,\mathbf{y}}\int d\mathbf{t}p^z_{n,\mathbf{t},\mathbf{y}}\sum_{k=1}^n\|D_x\Psi(x,\mathbf{t},\mathbf{y},t_k)\|\frac{\|D_x\lambda(z_{t_{t_k-}},y_{t_k})\|}{\lambda(z_{t_k-},y_{t_k})}$$
$$\leq M_\lambda^2\sum_{n,\mathbf{y}}\int d\mathbf{t}\sum_{k=1}^n\tilde p^k_{n,\mathbf{t},\mathbf{y}}\frac{V(z_{t_{k-}})}{\bar\lambda(z_{t_{k-}})}C_\gamma\exp(-\gamma t_k)$$
$$\leq M_\lambda^2 C_\gamma\sum_{k=1}^\infty\mathbb{E}^k\left(1_{\tau_k\leq T}\frac{V(Z_{\tau_k-})}{\bar\lambda(Z_{\tau_{k-}})}\right)$$
$$=M_\lambda^2 C_\gamma\mathbb{E}^z\sum_{k=1}^{N_T}\frac{V(Z_{\tau_{k-}})}{\bar\lambda(Z_{\tau_{k-}})}$$

where at the third step we notice that (Z_{τ_k-},τ_k) have the same law under $\mathbb{P}^z$ and $\mathbb{P}^k$. Also notice that $V(Z_{s-})/\bar\lambda(Z_{s-})$ is left continuous. On the other hand, by Proposition 26.7 of [10], the compensator of N_t is

$\int_0^t \bar{\lambda}(Z_{s-})ds$. Hence applying formula 31.18 of [10] with $b(Z_{\tau_k}, Z_{\tau_k-}) = V(Z_{\tau_k}-)/\bar{\lambda}(Z_{\tau_k}-)$ and $\delta = 0$, we have

$$\mathbb{E}^z \sum_{k=1}^{N_T} \frac{V(Z_{\tau_{k-}})}{\bar{\lambda}(Z_{\tau_{k-}})} = \mathbb{E}^z \int_0^T V(Z_{s-})ds = \int_0^T \mathbb{E}^z V(Z_s)ds \leq \frac{1}{\gamma}(V(z) + k_v T).$$

where we used the fact that Z_s jumps only countably many times and used Lemma 3.7 for the upper bound. □

Remark 4.4. *In* [7, 4]*, the contracting distance is set up through a concrete coupling mechanism. Usually this setup requires us to manage two copies of the process at the same time, which may be difficult in certain situations. On the other hand, the framework in* [18] *presented here requires only a perturbation analysis on one process. Thus the proof here is more straightforward, and the bounds used here have room to work for weaker conditions. Yet, the reason we can do such analysis is that a PDMP has a differential flow* Ψ*, which* [7] *does not assume.*

Remark 4.5. *The construction of the processes* $\widetilde{Z}_t^k$ *is not necessary. The bound for second line of* (4.8) *can be obtained through an advanced application of formula 31.18 of* [10] *with* $b(Z_{\tau_k}, Z_{\tau_k-}) = V(Z_{\tau_{k-}})/\lambda(Z_{\tau_k-}, Y_{\tau_k})$*. We are constructing these auxiliary processes to offer better probabilistic intuition.*

4.3 Accessibility and small sets

The verification of the small set condition in Theorem 4.1 is a standard one based on an accessibility study. We will also introduce a few notions for accessibility, as they will become useful when verifying Assumption 3.9 for the stochastic lattice models. Most of the following lemmas have a variant in either [3] or [7]. Since we are working with unbounded rates, and we want to keep this article self-contained, we are presenting these standard verifications as well.

First of all, the accessibility of process Z_t can be characterized by the density $p^{z,t}_{n,\mathbf{t},\mathbf{y}}$.

Definition 4.6. *A jump sequence* $(n, \mathbf{t}, \mathbf{y})$ *or simply written as* $(\mathbf{t}, \mathbf{y})$ *is admissible from* z *up to time* t*, if* $t_n < t$ *and* $p^{z,t}_{n,\mathbf{t},\mathbf{y}} > 0$*. A state* $z' \in E$ *or* $y' \in F$ *is called accessible from* z *at time* t *if there is an admissible jump sequence* $(n, \mathbf{t}, \mathbf{y})$ *from* z *up to time* t*, with* z_t *generated by* (4.1) *is* z' *or its* Y *part is* y'*.*

The following properties are immediate by the formula (4.2).

Lemma 4.7. *For a PDMP, the following holds:*

(1) $p^{(x,y),t}_{n,\mathbf{t},\mathbf{y}}$ *is continuous with respect to* $\mathbf{t} = (t_1, \ldots, t_n)$ *and* x *(hence aslo* z*);*

(2) *If* $(n,\mathbf{t},\mathbf{y})$ *and* $(n',\mathbf{t}',\mathbf{y}')$ *are admissible from* z *up to time* t *and* z_t *up to time* t'*, then* $p^{z,t+t'}_{n+n',\mathbf{t}+\mathbf{t}',\mathbf{y}+\mathbf{y}'} = p^{z,t}_{n,\mathbf{t},\mathbf{y}} p^{z_t,t'}_{n',\mathbf{t}',\mathbf{y}'}$*, where* $z_t = (\Psi(x,\mathbf{t},\mathbf{y},t), y_t)$ *and*

$$\mathbf{t}+\mathbf{t}' = (t_1,\dots,t_n,t+t'_1,\dots,t+t'_{n'}),\quad \mathbf{y}+\mathbf{y}' = (y_1,\dots,y_n,y'_1,\dots,y'_{n'}).$$

(3) *If* $p^{z,t}_{n,\mathbf{t},\mathbf{y}} > 0$ *then* $p^{z,t+s}_{n,\mathbf{t},\mathbf{y}} > 0$ *for* $s \geq 0$.

Proof. The first two claims are immediate based on formula (4.2). To see the third, observe that the trivial jump sequence $(0,\emptyset,\emptyset)$ is admissible from all z' up to any time s, as

$$p^{z',s}_0 = \exp\Big(-\int_0^s \bar{\lambda}(\Psi^{y'}_r x', y')dr\Big) \geq \exp\Big(-M_\lambda \int_0^s V(\Psi^{y'}_r x', y')dr\Big),$$

as Ψ is non-explosive and V is continuous in x, we find that $p^{z',s}_0 > 0$. Then by claim (2), $p^{z,t+s}_{n,\mathbf{t},\mathbf{y}} = p^{z,t}_{n,\mathbf{t},\mathbf{y}} p^{z_t,s}_0$ is strictly positive. □

The following lemma provides a straightforward way to verify Assumption 3.9 using the density.

Lemma 4.8. *With Assumption 3.6,*

(1) *If* $j \in F$ *is accessible from* $z \in E$ *at time* s*, then* $\mathbb{P}^z(Y_t = j) > 0$ *for any* $t \geq s$*;*

(2) *If Assumption 3.9 holds in addition, then for any compact set* C*, there exists constants* $t_0, m_0 > 0$ *such that*

$$\mathbb{P}^z(Y_{t_0} = y_c) \geq m_0, \quad \forall z \in C.$$

Proof. For the first claim, using Lemma 4.7 claim (3), we can find an admissible sequence $(n,\mathbf{t},\mathbf{y})$, i.e. $p^{z,t}_{n,\mathbf{t},\mathbf{y}} > 0$. Then by the continuity of $p^{z,t}_{n,\mathbf{t},\mathbf{y}}$ in $\mathbf{t}$ from Lemma 4.7 claim (1), there exists a neighbor $O_\mathbf{t}$ of $\mathbf{t}$ in the set $\{(t'_1,\dots,t'_n), t'_n < t\}$ such that if $\mathbf{t}' \in O_\mathbf{t}$, $p^{z,t}_{n,\mathbf{t},\mathbf{y}'} > 0$. Hence

$$\mathbb{P}^z(Y_t = j) \geq \mathbb{P}(Y_{s\leq t} \text{ goes through } \mathbf{y} \text{ with jump times in } O_\mathbf{t}) = \int_{O_\mathbf{t}} d\mathbf{t}' p^{z,t}_{n,\mathbf{t}',\mathbf{y}} > 0.$$

To see the second claim, by Assumption 3.9 and the expectation formula (4.3), for each $z \in C$, there exists an admissible sequence $(n,\mathbf{t},\mathbf{y})$ up to t such that $p^{z,t}_{n,\mathbf{t},\mathbf{y}} > 0$. As $p^{z,t}_{n,\mathbf{t},\mathbf{y}}$ is continuous in z, we can find a finite cover of C such that

$$p^{z,t_i}_{n,\mathbf{t},\mathbf{y}} > 0, \quad \forall z \in O_i.$$

Let $t_0 = \max\{t_i\}$, we have $\mathbb{P}^z(Y_{t_0} = y_c) > 0$ for all z by Lemma 4.7 claim (3) and the first claim of this lemma. Then using the compactness and the continuity in z again we find a uniform lower bound m_0 for the density. □

By the definition of small sets and the construction of d, (4.5), we prove the following stronger claim to verify the small set condition of Theorem 4.1.

Lemma 4.9. *Under the conditions of Theorem 3.10, for any fixed strictly positive M,ε, there exists strictly positive constants t_1 and m_1, such that for any $z,z'\in E$ satisfying $V(z),V(z')\le M$, there exists a coupling of $\mathbb{P}^z_{t_1},\mathbb{P}^{z'}_{t_1}$ that satisfies:*

$$\mathbb{P}^{z,z'}(d(Z_{t_1},Z'_{t_1})\le 2\varepsilon)\ge m_1.$$

Hence $\{z:V(z)\le M\}$ is d-small.

Proof. As d is a distance, by Lemma A.1, it suffices for us to show that

$$\mathbb{P}^z(d(Z_{t_1},z_c)\le\varepsilon)\ge m_1,\quad \forall z:V(z)\le M.$$

Where $z_c=(y_c,x_c)$ is given by Assumption 3.9. The proof proceeds through two steps: we first couple Y_t,Y'_t to y_c. Then we keep the value of Y to be y_c afterwards, until the contracting dynamics brings the X part close enough. In detail, by Lemma 4.8 (2), there is a constant t_0 and c_0, such that when $V(z)\le M$, $\mathbb{P}^z(Y_{t_0}=y_c)\ge m_0$. On the other hand, as $V(z)\ge\|x\|^2$, by the Markov inequality, if we let $d_v=\sqrt{2K_{t_0}M/m_0}$ with K_{t_0} given by Lemma 3.7,

$$\mathbb{P}^z(\|X_{t_0}\|\ge d_v)\le\frac{\mathbb{E}^zV(Z_{t_0})}{d_v^2}\le\frac{K_{t_0}M}{d_v^2}\le\frac{1}{2}m_0.$$

Hence by the lower bound for intersection of events:

$$\mathbb{P}^z(Y_{t_0}=y_c,\|X_{t_0}\|\le d_v)\ge\frac{1}{2}m_0,\quad \forall z:V(z)\le M.$$

As the attractor x_c satisfies $V(x_c,y_c)<\infty$ and V is continuous in the x part, we can find $\bar{M}_\lambda$ such that $V(u,y_c)\le\bar{M}_\lambda$ when u is inside the following set:

$$\{\Psi^{y_c}_t x:\|x\|\le d_v,t\ge 0\}\cup\{x+s(x_c-x):s\in[0,1],\|x\|\le 1+\|x_c\|\}. \tag{4.9}$$

Hence, if we let $\varepsilon'\le\delta\bar{M}_\lambda^{-1}\varepsilon$, then for any x' such that $\|x'-x_c\|\le\varepsilon'$,

$$d((x',y_c),(x_c,y_c))\le\delta^{-1}\|x'-x_c\|\int_0^1 V(x'+s(x_c-x'),y_c)ds\le\varepsilon'\delta^{-1}\bar{M}_\lambda\le\varepsilon.$$

Pick a T' such that $C_\gamma(d_v+\|x_c\|)\exp(-\gamma T')\le\varepsilon'$. Then by Assumption 3.8, for $\|x\|\le d_v$:

$$\begin{aligned}\|\Psi^{y_c}_{T'}x-x_c\| &= \|\Psi^{y_c}_{T'}x-\Psi^{y_c}_{T'}x_c\|\\ &\le\|x-x_c\|\int_0^1\|D_x\Psi^{y_c}_{T'}(x+r(x_c-x))\|dr\le\varepsilon'.\end{aligned}$$

Thus, we can generate the following bound of probability using $V(u,y_c)\le\bar{M}_\lambda$ for $u\in$ (4.9) and applying formula (4.2) with $n=0$ and the Markov property:

$$\begin{aligned}\mathbb{P}^z(d(Z_{t_0+T'},z_c)\leq\varepsilon) &\geq \mathbb{P}^z(\|X_{t_0}\|\leq d_v, Y_s=y_c, s\in[t_0,t_0+T'])\\ &\geq \mathbb{E}^z\left[1_{\|X_{t_0}\|\leq d_v, Y_{t_0}=y_c}\exp\left(-\int_0^{T'}\bar{\lambda}(\Psi_s^{y_c}X_{t_0},y_c)ds\right)\right]\\ &\geq \frac{1}{2}m_0\exp(-\bar{M}_\lambda T') =: m_1.\end{aligned}$$

This concludes our first claim with $t_1=t_0+T'$. To see this actually implies that $\{V(z)\leq M\}$ is d-small, simply let $2\varepsilon=1/2$, with the T given by Lemma 4.3, we have the following for any $V(z),V(z')\leq M$,

$$\begin{aligned}d(\mathbb{P}^z_{t_1+T},\mathbb{P}^{z'}_{t_1+T}) &\leq \int \mathbb{P}^{z,z'}(Z_{t_1}\in dw, Z'_{t_1}\in dw')d(\mathbb{P}^w_T,\mathbb{P}^{w'}_T)\\ &\leq \mathbb{P}^{z,z'}(d(Z_{t_1},Z'_{t_1})\geq 1/2)+\frac{1}{2}\mathbb{P}^{z,z'}(d(Z_{t_1},Z'_{t_1})\leq 1/2)\leq 1-\frac{1}{4}m_1.\end{aligned}$$

□

4.4 Proof of Theorem 3.10

With the conditions of Theorem 4.1 verified, it is rather elementary to show Theorem 3.10.

Proof. Using the Lyapunov function, one can bound $\mathbb{E}\|X_t\|^2$ uniformly in t. This indicates that the family of measures $\{\mathbb{P}^z_t\}_{t\geq 0}$ is uniformly tight, so by Krylov-Bogolyubov theorem [8] there exists at least one invariant measure π. To see the geometric convergence in m, based on Lemma 4.2, 4.3 and 4.9, Theorem 4.1 can be applied to Z_t with metric defined by (4.5). With the metric $\tilde{d}$ defined by Theorem 4.1, $\tilde{d}(\mathbb{P}^z_t,\mathbb{P}^{z'}_t)\leq\frac{1}{2}\tilde{d}(z,z')$. By recursively applying this relation with the Markov property, i.e. $\mathbb{P}^z_{nt}=P^*_t\mathbb{P}^{z'}_{(n-1)t}$

$$\tilde{d}(\mathbb{P}^z_{nt},\mathbb{P}^{z'}_{nt})\leq\frac{1}{2}\tilde{d}(\mathbb{P}^z_{(n-1)t},\mathbb{P}^{z'}_{(n-1)t})\leq\cdots\leq\frac{1}{2^n}\tilde{d}(z,z')\leq\frac{1}{2^n}\sqrt{1+V(z)+V(z')}$$

Recall the convention that $V(z)\geq 1$, so for sufficiently large n, if event

$$A=\left\{\tilde{d}(Z_{nt},Z'_{nt})\leq\frac{\sqrt{1+V(z)+V(z')}}{\sqrt{2}^n}\right\}$$

takes place, $d(Z_{nt},Z'_{nt})\leq\sqrt{1+V(z)+V(z')}/2^n$, which implies $Y_{nt}=Y'_{nt}$ and $m(Z_{nt},Z'_{nt})\leq\delta d(Z_{nt},Z'_{nt})$, therefore

$$m(Z_{nt},Z'_{nt})\leq\delta d(Z_{nt},Z'_{nt})\leq\frac{\delta\sqrt{1+V(z)+V(z')}}{2^n}.$$

On the other hand, by the Markov inequality

$$\mathbb{P}(A^c)\leq\frac{\tilde{d}(\mathbb{P}^z_{nt},\mathbb{P}^{z'}_{nt})}{1/\sqrt{2^n}}\leq\frac{1}{\sqrt{2^n}}.$$

Therefore by the Cauchy inequality

$$
\begin{aligned}
m(\mathbb{P}^{z}_{nt},\mathbb{P}^{z'}_{nt}) &= \mathbb{E}1_A m(Z_{nt},Z'_{nt}) + \mathbb{E}1_{A^c} m(Z_{nt},Z'_{nt})\\
&\le \frac{\delta}{2^n}\sqrt{1+V(z)+V(z')} + [\mathbb{P}(A^c)\mathbb{E}(m^2(Z_{nt},Z'_{nt}))]^{\frac{1}{2}}\\
&\le \frac{\delta}{2^n}\sqrt{1+V(z)+V(z')} + \left[\frac{1}{2}\sqrt{2^n}\mathbb{E}(1+2\|Z_{nt}\|^2+2\|Z'_{nt}\|^2)\right]^{\frac{1}{2}}\\
&\le \frac{\delta}{2^n}\sqrt{1+V(z)+V(z')} + \left[\frac{1}{\sqrt{2^n}}(2V(z)+2V(z')+2k_v/\gamma+1)\right]^{\frac{1}{2}}
\end{aligned}
$$

where in the last step we used $\|x\|^2 \le V(z)$ and Lemma 3.7. In view of such bound, it is clear the last claim of this theorem is proved. To see this implies exponential convergence from μ to π, it suffices go through the following standard procedure:

$$
\begin{aligned}
m(\mathbb{P}^{\mu}_t,\pi) = m(\mathbb{P}^{\mu}_t,\mathbb{P}^{\pi}_t) &\le \int \mu(dz)\pi(dz')\mathbb{E}^{z,z'}m(Z_t,Z'_t)\\
&= \int \mu(dz)\pi(dz')m(\mathbb{P}^{z}_t,\mathbb{P}^{z'}_t)\\
&\le C\exp(-\beta t)\int \mu(dz)\pi(dz')\sqrt{V(z)+V(z')+1}.
\end{aligned}
$$

□

5 Geometric ergodicity for the simplest tropical stochastic climate model

In this section, we prove Theorem 2.1 by applying Theorem 3.10 to the simplest tropical climate model introduced in Section 2.1. Recall its formulation:

(5.1)

$$
\begin{aligned}
\frac{dK_i}{dt} + D_x^+ K_i &= -\frac{\bar{d}+d_\theta+d_{sh}}{2}K_i - \frac{d_\theta+d_{sh}-\bar{d}}{2}R_i - (d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i}+P_i),\\
\frac{dR_i}{dt} - D_x^- R_i &= -\frac{\bar{d}+d_\theta+d_{sh}}{2}R_i - \frac{d_\theta+d_{sh}-\bar{d}}{2}K_i - (d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i}+P_i),\\
\frac{dZ_i}{dt} &= -d_qZ_i + \frac{d_\theta+d_{sh}-d_q}{2}\bar{Q}(K_i+R_i) + \bar{Q}(d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i}) + d_qq_{s,i} - (1-\bar{Q})P_i,\\
P_i &= (1-\sigma_I)\tau_c^{-1}(Z_i + \frac{\alpha+\bar{Q}}{2}(K_i+R_i)-\hat{q})^+,\\
c_a(\eta_i) &= \frac{l-\eta_i}{\tau_I},\quad c_d(\eta_i) = \frac{\eta_i}{\tau_I}\exp\bigg(-2U_0\frac{\eta_i-1}{l-1}+\gamma(Z_i+\frac{1}{2}\bar{Q}(K_i+R_i))-h_0\bigg).
\end{aligned}
$$

Using the terminology of Section 3, the ODE part is

$$
\mathbf{U}_t = (K_i(t),R_i(t),Z_i(t))_{i\in I} \in \mathscr{H} = \mathbb{R}^{3I},
$$

and the Markov jump part is $\eta_t = (\eta_i(t))_{i\in I} \in F = \{0,\dots,l\}^I$. Since conditioned on value of $\mathbf{U}_t$, $\eta_i(t)$ are independent birth/death processes, the transition rates of η_t are as follow:

$$\begin{aligned}
&\lambda(U,\eta,\eta+e_i) = c_a(\eta_i) = \frac{l-\eta_i}{\tau_I} \geq \frac{1}{\tau_I} > 0, \quad \eta_i \leq l-1 \\
&\lambda(U,\eta,\eta-e_i) = c_d(U_i,\eta_i) \geq \frac{c_0}{\tau_I} > 0, \quad \eta_i \geq 1, \\
&\lambda(U,\eta,\eta') = 0, \quad \forall \eta' \neq \eta \pm e_i, \quad \text{or} \quad \eta' \notin F,
\end{aligned} \tag{5.2}$$

where we used the notation $\eta \pm e_i := (\eta_1,\dots,\eta_i \pm 1,\dots,\eta_l)$. In other words, a transition from η is possible if and only if it is toward a neighboring site of η on $F = \{0,\dots,l\}^I$, regardless of the state of X_t. Hence the first item of Assumption 3.6 holds.

5.1 Path-wise dissipative energy

As the Lyapunov function V plays the key role in all parts of our theory, let us find it first. Under the condition of Theorem 2.1, system (5.1) has a notion of energy that is path-wise dissipative, which is motivated from a continuum version in [15, 32].This is much stronger than the requirement of Assumption 3.4, as it implies a compact invariant set for the dynamics. Consider the following energy at site $i \in I$,

$$\varepsilon_i = \frac{1}{2}(K_i^2 + R_i^2) + \frac{Z_i^2}{(\alpha+\bar{Q})(1-\bar{Q})}. \tag{5.3}$$

Lemma B.1 produces the following dissipative principle with a proper constant $\gamma > 0$ and a function k_v:

$$\frac{d\varepsilon_i}{dt} \leq -K_i D_x^+ K_i + R_i D_x^- R_i - \gamma\varepsilon_i + k_v(\theta_{s,i},\theta_{eq,i},q_{s,i}).$$

To see why this implies path-wise dissipation, we need the following property of the operator $D_x^{\pm}$, which is one of reasons why the upwind discretization scheme stabilizes a PDE:

Lemma 5.1. *For any $(f_i)_{i=1,\dots,N}$, the following holds:*

$$\sum_{i=0}^{N-1} f_i D_x^+ f_i \geq 0, \quad \sum_{i=0}^{N-1} f_i D_x^- f_i \leq 0.$$

Proof. The sums can be written as:

$$\sum_{i=0}^{N-1} f_i D_x^+ f_i = N^{-1}\Big(\sum_{i=0}^{N-1} f_i^2 - \sum_{i=0}^{N-1} f_i f_{i-1}\Big),$$

$$\sum_{i=0}^{N-1} f_i D_x^- f_i = N^{-1}\Big(\sum_{i=0}^{N-1} f_i f_{i+1} - \sum_{i=0}^{N-1} f_i^2\Big).$$

Our claim follows from the Young's inequality: $f_i f_{i-1} \leq \frac{1}{2}(f_i^2 + f_{i-1}^2)$. $\square$

Hence if we let $V = \sum_{i\in I} \varepsilon_i = V(x)$, it is path-wise dissipative:

$$\frac{dV}{dt} = \sum_{i\in I} \frac{d\varepsilon_i}{dt} \le -\gamma V + \sum_i k_v(\theta_{s,i}, \theta_{eq,i}, q_{s,i}).$$

This implies V (or $V+1$ if one wants $V \ge 1$), satisfies the requirement of Assumption 3.4. Moreover, this induces that $K = \{(x,\eta) : V(x) \le \sum_i k_v(\theta_{s,i}, \theta_{eq,i}, q_{s,i})/\gamma\}$ is actually an absorbing invariant set; in other words, for any $z \in E$, there exists a $t_0(z)$ such that $Z_t \in K$ for $t \ge t_0(z)$, $\mathbb{P}^z$-a.s, and $t_0(z) = 0$ if $z \in K$. So in the long term, we can assume (5.1) actually takes place in K. Since $F = \{0,\dots,l\}^I$ is finite, K is compact. Since the transition rates $c_a(\eta_i)$ and $c_d(\eta_i)$ are smooth with respect to X_i, restraining the process to be in K provides a trivial upper bounds for the transition rates and their Frechét derivatives. In other words, Assumption 3.6 holds for the simplest tropical climate model given by (5.1).

5.2 Piecewise contraction

The piecewise contraction condition, Assumption 3.8, can usually be verified through analyzing the propagation of a perturbation in the initial condition. Indeed, for any $h \in \mathscr{H}$, let $X_t^h := \langle D_x\Psi(x, Y_{s\le t}, t), h\rangle$, which is the perturbation on X_t caused by a perturbation on X_0 in the direction of h. According to the differential flow formulation of Ψ, we have:

$$\begin{aligned} X_t^h &= \langle D_x X_0, h\rangle + \int_0^t \langle D_x \psi(\Psi(x, Y_{r\le s}, s), Y_s), h\rangle ds \\ &= h + \int_0^t \nabla_x \psi(Z_s) \cdot \langle D_x\Psi(x, Y_{r\le s}, s), h\rangle ds \\ &= h + \int_0^t \nabla_x \psi(Z_s) \cdot X_s^h ds. \end{aligned}$$

In other words, X_t^h is the solution to the ODE:

$$\frac{dX_t^h}{dt} = \nabla_x \psi(Z_t) \cdot X_t^h, \quad X_0^h = h. \tag{5.4}$$

This is also known as the derivative flow for process X_t. Then in order to verify $\|D_x\Psi(x, Y_{s\le t}, t)\| \le C_\gamma e^{-\gamma t}$, it suffices to show

$$\|X_t^h\| \le C_\gamma e^{-\gamma t}, \quad \forall h \in \mathscr{H}, \|h\| \le 1.$$

This verification method is advantageous for PDMP with simple differential flows, as the transition rates are not relevant in the formula.

Applying this method to (5.1), while denoting $f^h = \langle D_x f, h\rangle$ for any variable f, the derivative flow of (5.1) is then

(5.5)

$$\begin{aligned}
\frac{dK_i^h}{dt} + D_x^+ K_i^h &= -\frac{\bar{d} + d_\theta + d_{sh}}{2} K_i^h - \frac{d_\theta + d_{sh} - \bar{d}}{2} R_i^h - P_i^h, \quad K_i^h(0) = h_{K,i},\\
\frac{dR_i^h}{dt} - D_x^- R_i^h &= -\frac{\bar{d} + d_\theta + d_{sh}}{2} R_i^h - \frac{d_\theta + d_{sh} - \bar{d}}{2} K_i^h - P_i^h, \quad R_i^h(0) = h_{R,i},\\
\frac{dQ_i^h}{dt} &= -d_q Q_i^h + \frac{d_\theta + d_{sh} - d_q}{2} \bar{Q}(K_i^h + R_i^h) - (1-\bar{Q})P_i^h, \quad Q_i^h(0) = h_{Q,i}.
\end{aligned}$$

This is essentially a homogeneous version of (2.2). One sufficient condition to show path-wise contraction is through showing for a $\gamma > 0$:

$$\sum_{i\in I} |K_i^h|^2 + |R_i^h|^2 + |Z_i^h|^2 \le |h|^2 \exp(-\gamma t).$$

Inspired by the construction of Lyapunov function in Section 5.1, it is intuitive to consider the following quantity:

$$\varepsilon_i^h = \frac{1}{2}\Big((K_i^h)^2 + (R_i^h)^2\Big) + \frac{(Q_i^h)^2}{(1-\bar{Q})(\alpha+\bar{Q})}. \tag{5.6}$$

By Lemma B.2, there is a $\gamma > 0$, such that

$$\frac{d\varepsilon_i^h}{dt} \le -K_i^h D_x^+ K_i^h + R_i^h D_x^- R_i^h - \gamma \varepsilon_i^h.$$

So by Lemma 5.1, $V_h = \sum_i \varepsilon_i^h$ is exponentially decaying:

$$\frac{dV_h}{dt} \le -\gamma V_h.$$

This implies $\sum_i |K_i^h|^2 + |R_i^h|^2 + |Z_i^h|^2$ decay exponentially in time, as V_h dominates this norm. Therefore the model (5.1) satisfies Assumption 3.8.

5.3 Accessibility

Through the discussion of Section 5.1, there is a compact attracting invariant set, in which the Lyapunov function is bounded from above. This implies every state $\eta \in F$ has its attractor $\mathbf{U}_\eta$ which satisfies $V(U_\eta, \eta) < \infty$. As the transition is possible when two states are neighbors, and the finite set $F = \{1, \dots, l\}^I$ is connected through neighboring relations, so for any two state $\eta, \eta' \in F$, we can find a path $\vec{\eta} = (\eta^1, \dots, \eta^n = \eta')$ such that η^1 and η, η^k and η^{k+1} are neighbors. Then it is elementary to verify that $P^{z,t}_{n,\mathbf{t},\vec{\eta}} > 0$ for any $\mathbf{t} = (t_1, \dots, t_n)$ with $t_1 < \cdots < t_n$, since V is bounded. So by Lemma 4.8, Assumption 3.9 is verified.

5.4 Concluding Remarks

As the Assumptions 3.4, 3.6, 3.8 and 3.9 are verified, Theorem 3.10 can directly apply to the simplest tropical climate model given by (5.1). One of key features here is the energy V is path-wise dissipative. Using it, we actually find a compact invariant set, which enables Assumption 3.6 to hold rather trivially. In fact, this also makes the results of [7, 4] directly applicable to system (5.1). On the other hand, if such a compact set cannot be found, the verification maybe much more difficult, as we will soon find out in the next section.

6 Geometric Ergodicity for the stochastic skeleton model

In this section, we prove Theorem 2.3 by applying Theorem 3.10 to the skeleton model for MJO [40, 39] introduced in Section 2.2. Recall its formulation:

$$\begin{aligned}
&\frac{dK_i}{dt}+D_x^+K_i=(S_i^\theta-\bar{H}A_i)/2-\bar{d}K_i,\\
&\frac{dR_i}{dt}-D_x^-R_i/3=(S_i^\theta-\bar{H}A_i)/3-\bar{d}R_i,\\
&\frac{dZ_i}{dt}=(S_i^\theta-\bar{H}A_i)(1-\bar{Q})-\bar{d}Z_i,\\
&Q_i=Z_i+\bar{Q}(K_i+R_i),\quad A_i=\Delta A\eta_i\\
&c_a(\eta_i,Q_i)=\begin{cases}\Gamma\gamma'|Q_i|\eta_i+1_{\eta_i=0} & Q_i\geq 0\\ 1_{\eta_i=0} & Q_i<0\end{cases},\\
&c_d(\eta_i,Q_i)=\begin{cases}0 & Q_i\geq 0\\ \Gamma\gamma'|Q_i|\eta_i & Q_i<0\end{cases}.
\end{aligned}\tag{6.1}$$

Using the terminology of Section 3, the ODE part is

$$\mathbf{U}_t=(K_i(t),R_i(t),Z_i(t))_{i\in I}\in\mathscr{H}=\mathbb{R}^{3I},$$

and the Markov jump process part is $\eta_t=(\eta_i(t))_{i\in I}\in F=\mathbb{N}^I$. The transition rate of η_t has the same formulation as in (5.2), but with c_a and c_d given by (6.1).

6.1 Lyapunov Structure

Motivated by the continuum energy conservation principle developed in [34, 40], system (6.1) also has a dissipative energy:

$$\mathscr{E}:=\sum_i\frac{1}{2}\left[2K_i^2+3R_i^2+\frac{(Z_i+1)^2}{(1-\bar{Q})\bar{Q}}\right]+\frac{\bar{H}\Delta A\eta_i}{\Gamma\bar{Q}}+1.\tag{6.2}$$

Combining the results of Lemma B.3 and 5.1, there exists a $\gamma > 0$ such that,

$$\mathscr{L}\mathscr{E} \le -\gamma\mathscr{E} + \sum_{i\in I} k_v(S_i^\theta). \tag{6.3}$$

Unfortunately, $\mathscr{E}$ does not satisfies Assumption 3.6, since $\bar{\lambda}$ is roughly $\sum |Q_i|\eta_i$, which is not bounded by $\mathscr{E}$. So instead, we will use its cubic, $\mathscr{E}^3$, to be our Lyapunov function. $\mathscr{E}^3$ satisfies Assumption 3.4 because of the following lemma:

Lemma 6.1. *Assume that a function $\mathscr{E} : E \mapsto \mathbb{R}_+$ is a Lyapunov function with jumps $\Delta\mathscr{E}_t = \mathscr{E}(X_t, Y_t) - \mathscr{E}(X_t, Y_{t-})$ bounded by a constant B, and the total jump intensity of the PDMP is bounded by $\bar{\lambda}(z) \le M_\lambda \mathscr{E}^{2-\alpha}(z)$, $\alpha >$ 0. Then for any $n \in \mathbb{N}$, $n \ge 1$, $V = \mathscr{E}^n$ is also a Lyapunov function.*

Proof. Recall the formal generator $\mathscr{L}$ for $\mathscr{E}$ is:

$$\mathscr{L}\mathscr{E} = \psi(z)\partial_x\mathscr{E}(z) + \sum_{y'\in F} \lambda(z,y')(\mathscr{E}(x,y') - \mathscr{E}(x,y)).$$

Hence

$$\begin{aligned}\mathscr{L}\mathscr{E}^n &= n\mathscr{E}^{n-1}(z)\psi(z)\partial_x\mathscr{E}(z) + \sum_{y'\in F} \lambda(z,y')(\mathscr{E}^n(x,y') - \mathscr{E}^n(x,y)) \\ &= \sum_{y'\in F} \lambda(z,y')(\mathscr{E}(x,y') - \mathscr{E}(z)) \sum_{k=0}^{n-1} (\mathscr{E}^{n-1-k}(x,y')\mathscr{E}^k(z) - \mathscr{E}^{n-1}(z)) \\ &\quad + n\mathscr{E}^{n-1}(z)\mathscr{L}\mathscr{E}(z).\end{aligned}$$

By the bounded jumps condition, $|\mathscr{E}(x,y') - \mathscr{E}(z)| \le B$ for all y', $\lambda(z,y') >$ 0, so with some constant D the following holds:

$$\begin{aligned}(\mathscr{E}(x,y') - \mathscr{E}(z))(\mathscr{E}^{n-1-k}(x,y')\mathscr{E}^k(z) - \mathscr{E}^{n-1}(z)) &= 0, \quad k = n-1; \\ (\mathscr{E}(x,y') - \mathscr{E}(z))(\mathscr{E}^{n-1-k}(x,y')\mathscr{E}^k(z) - \mathscr{E}^{n-1}(z)) &\le D\mathscr{E}^{n-2}(z), \quad \forall k \le n-2.\end{aligned}$$

Therefore, by Young's inequality, $\sum_{y'} \lambda(z,y') \le M_\lambda \mathscr{E}^{2-\alpha}$ and the fact that $\mathscr{E}$ satisfies Assumption 3.4 there is a $\tilde{k}_v$ such that the following holds:

$$\mathscr{L}\mathscr{E}^n \le -n\gamma\mathscr{E}^n(z) + nk_v\mathscr{E}^{n-1}(z) + nD\mathscr{E}^{n-\alpha}(z) \le -\frac{1}{2}\gamma\mathscr{E}^n(z) + \tilde{k}_v.$$

The other requirements of Lyapunov function in Assumption 3.4 can be easily verified by $\mathscr{E}$ being a Lyapunov function. $\square$

With Lemma 6.1, $V = \mathscr{E}^3$ will be a proper choice of Lyapunov function for system (6.1), moreover Assumption 3.6 is also satisfied with this choice:

Lemma 6.2. *Assumptions 3.4 and 3.6 are satisfied by model* (6.1) *with $V = \mathscr{E}^3$ being the Lyapunov function, where $\mathscr{E}$ is given by* (6.2).

Proof. The total jump intensity satisfies:

$$\begin{aligned}\bar{\lambda}(X,\eta) &= \sum_{i\in I} c_a(\eta_i, Q_i) + c_d(\eta_i, Q_i)\\ &= \sum_{i\in I}(1_{\eta_i=0} + \Gamma|Q_i|\eta_i).\end{aligned}$$

Moreover, the total number of possible transitions sites from any η is at most $2N+1$. On the other hand, recall that $Q_i = \bar{Q}(K_i+R_i)+Z_i$, so using $\eta_i \geq 0$ and Young's inequality, for some c_1, c_2, c_3 the following holds:

$$\begin{aligned}\mathscr{E}^3 &\geq \sum_i c_1\left(2K_i^2 + 3R_i^2 + \frac{(Z_i+1)^2}{(1-\bar{Q})\bar{Q}}\right)\left(\frac{\bar{H}\Delta A\eta_i}{\Gamma\bar{Q}}\right)^2 + c_1\left(\frac{\bar{H}\Delta A\eta_i}{\Gamma\bar{Q}}\right)^3 + 1\\ &\geq \sum_i c_2[\bar{Q}(K_i+R_i)+(Z_i+1)]^2\eta_i^2 + c_2\eta_i^3 + 1\\ &\geq \sum_i c_2 Q_i^2\eta_i^2 + 2c_2Q_i\eta_i^2 + c_2\eta_i^2 + c_3\eta_i^2 \geq \sum_i \frac{c_2c_3}{c_2+c_3}Q_i^2\eta_i^2.\end{aligned}$$

Therefore, by Cauchy-Schwartz, $\bar{\lambda} \leq M_\lambda \mathscr{E}^{3/2}$ for some M_λ. For the Frechét derivative of the transition rates in Assumption 3.6, observe that

$$\|D_x c_a(\eta_i, Q_i)\|, \|D_x c_d(\eta_i, Q_i)\| \leq \Gamma|\eta_i| \leq \Gamma\mathscr{E}, \quad \|D_x\bar{\lambda}(x,\eta)\| \leq \sum_i \Gamma|\eta_i| \leq \Gamma\mathscr{E}.$$

Combine this with $\bar{\lambda}(z) \leq M_\lambda \mathscr{E}^{3/2}$, we can further enlarge M_λ such that Assumption 3.6 holds for $V = \mathscr{E}^3$.

On the other hand, since $\mathscr{E}$ has jumps coming only from the jumps of η_i, the jumps of $\mathscr{E}$ are bounded by $\bar{H}\Delta A/\Gamma\bar{Q}$ in size, and $\bar{\lambda} \leq M_\lambda\mathscr{E}^{3/2}$. Hence it is easy to verify all the conditions of Lemma 6.1 for $\mathscr{E}$ using the bound (6.3), so $V = \mathscr{E}^3$ is a Lyapunov function. □

6.2 Piecewise contraction

It is relative easy to verify Assumption 3.8 for system (2.7) using the derivative flow method described in Section 5.2. Using the same notation as there, the propagation of perturbation by (5.4) follows:

$$\begin{aligned}\frac{dK_i^h}{dt} + D_x^+K_i^h &= -\bar{d}K_i^h, \quad K_i^h(0) = h_{K,i}\\ \frac{dR_i^h}{dt} - D_x^-R_i^h/3 &= -\bar{d}R_i^h, \quad R_i^h(0) = h_{R,i}\\ \frac{dZ_i^h}{dt} &= -\bar{d}Z_i^h, \quad Z_i^h(0) = h_{Z,i}.\end{aligned}$$

This is evidently dissipative, since if we let $\mathscr{E}_h = \frac{1}{2}\sum_i (K_i^h)^2 + (R_i^h)^2 + (Z_i^h)^2$, straight forward computation combined with Lemma 5.1 gives:

$$\partial_t\mathscr{E}_h = \sum_i(-K_i^hD_x^+K_i^h + \frac{1}{3}R_i^hD_x^-R_i^h) - [\bar{d}(K_i^h)^2 + \bar{d}(R_i^h)^2 + \bar{d}(Z_i^h)^2] \leq -\bar{d}\mathscr{E}_h.$$

Therefore, $\sum_i (K_i^h)^2 + (R_i^h)^2 + (Z_i^h)^2$ decays exponentially fast in time.

6.3 Accessibility Study

We will verify Assumption 3.9 through the following stronger claim:

Lemma 6.3. *Under the condition of Theorem 2.3, $\eta = \vec{1}$, i.e. $\eta_i = 1$ for all i, is a state that is accessible from any $z \in E$ for system* (2.7).

Due to the degeneracy of the transition rate, discussed in Section 2.2, the verification is highly nontrivial.

Intuition

Before we give out the proof of Lemma 6.3, let us first illustrate the intuition. We will essentially design a jump sequence to reach state $\vec{1}$. The basic components of this jump sequence is the following two:

- $\mathbf{U}_t$ is a continuous process, so in a sufficiently short time, it is possible to have any finite number of jumps in η_t without changing the value of $\mathbf{U}_t$ much, we will use the verb "burst" to describe such mechanics;
- Ψ is a contracting dynamics, so with η_t remains a constant, after a sufficiently long time, $\mathbf{U}_t$ will converge to the attractor for η_t. We use the verb "converge" to describe such mechanics. In fact, the attractor with given value of η_i can be written down explicitly, which will be Lemma 6.7.

We will design the mechanics of jumps through the following case by case study, most of them are based on the jump intensity of (6.1).

(1) If $\eta_i = 0$, then we can change it to 1;
(2) If $Q_i < 0$, then we can burst η_i to 0, then use item (1) to change it to 1;
(3) If $Q_i > 0$, then we can burst η_i to any large number, or equivalently $D_i = S_i^\theta - \bar{H}\Delta A\eta_i$ is sufficiently small while keeping other D_j the same. Then the attractor for the new set of D_i, based on Lemma 6.7, is a state where K_j, R_j are all sufficiently small, hence $Q_j < 0$ for all $j \in I$, so we can use case (2) at each lattice point to get to destination $\vec{1}$;
(4) If $Q_i = 0$ for all $i \in I$, then by the no fixed point condition of Theorem 2.3, after a sufficiently long time, $Q_i \neq 0$ for some i since $\mathbf{U}_t$ will converge to the attractor;
(5) In view of (2), (3) and (4), we only need to consider the case when some of the Q_i are negative while some are zero and show that we can escape from this scenario to the previous situations. As I is finite, there is a state among all states that is accessible while its number of $J = \{i : Q_i = 0\}$ is at the minimum, let us show there is actually no $Q_i = 0$ for this minimum state. Assume the opposite, then at this minimum state there is an i such that $Q_i = 0, Q_{i-1} < 0$. In the view of (2), we could burst η_{i-1} to 0 or 1. These alternative choices will generate the contradiction in a weak hypo-elliptic fashion. In both cases, there

exist a short period of time $[0,\delta]$, such that $J^c = \{i : Q_i < 0\}$ does not decrease as Q_i are continuous. Then by the minimum assumption, J will remain the same set, meaning $Q_i = 0$ for $i \in J$ as $t \in [0,\delta]$. This implies a delicate balance:

$$0 = \frac{d}{dt}Q_i = \bar{Q}(-D_x^+ K_i + D_x^- R_i/3) + D_i(1-5/6\bar{Q}) - \bar{Q}(\bar{d}K_i + \bar{d}R_i) - \bar{d}Z_i.$$

Taking the time derivative again, we have

$$0 = \frac{d^2}{dt^2}Q_i = \frac{1}{N}\bar{Q}[-\dot{K}_i + \dot{R}_i/3] + \frac{1}{N}\bar{Q}[\dot{K}_{i-1} - \dot{R}_{i+1}/3] - \bar{Q}(\bar{d}\dot{K}_i + \bar{d}\dot{R}_i) - \bar{d}\dot{Z}_i.$$

This equation contains the term η_{i-1} from the term $\dot{K}_{i-1}$, which has values either 0 or 1 depends on the burst choice. Yet the other terms should be relatively close for both burst choices. This leads to a contradiction.

Preparation

The following two Lemmas make the illustrations of the two jump mechanisms rigorous. They actually hold for general PDMP and piecewise contracting systems:

Lemma 6.4 (Burst Mechanism). *Let Z_t be a PDMP, suppose for some fixed $z_0 \in E$ and a sequence in F, $y_0, y_1, \dots, y_n$ such that*

$$\lambda(x_0, y_i, y_{i+1}) > 0, \quad i = 0,1,\dots,n-1.$$

Then for any $\varepsilon > 0$, there is a sequence of jumping times $\mathbf{t} = (t_1,\dots,t_n)$ such that $p^{z_0,t}_{n,\mathbf{t},\mathbf{y}} > 0$, while $t_n \le \varepsilon$ and

$$\|x_s - x\| \le \varepsilon \quad \text{with } x_s := \Psi(x,\mathbf{t},\mathbf{y},s), s \le t_n.$$

Proof. By Assumption 3.6, λ is continuous in x, we can find $\delta, M > 0$ such that the following holds:

$$\lambda(x,y_i,y_{i+1}) > 0, \quad \forall \|x - x_0\| \le \delta, i = 0,1,\dots,n-1,$$
$$M := \sup\{\|\psi(x,y_i)\|, \|x-x_0\| \le \delta, i = 0,\dots,n\} < \infty.$$

Let

$$\xi := \frac{1}{n+1}(\varepsilon \wedge \frac{\delta}{M}); \qquad t_k = \xi k, \quad k = 1,\dots,n.$$

Then as $x_t = x_{t_{k-1}} + \int_{t_{k-1}}^t \psi(x_s, y_{k-1})ds$ for $t \in [t_{k-1}, t_k]$,

$$\|x_t - x_{t_{k-1}}\| \le M\xi \le \frac{\delta}{n} \quad t \in [t_{k-1}, t_k];$$

hence $\|x_t - x_0\| \le \delta$ for $t \in [0,t_n]$. By the setting of δ,

$$p^{z_0,\varepsilon}_{n,\mathbf{t},\mathbf{y}} = \exp\left(-\int_0^\varepsilon \bar{\lambda}(z_s)ds\right)\prod_{k=1}^n \lambda(x_{t_k}, y_{k-1}, y_k) > 0.$$

□

Lemma 6.5 (Contraction mechanism). *With Assumption 3.8, then for any open set $O\subset E, y\in F$, suppose the attractor for y is in O, then for any $x\in E$, there exists an T such that $\Psi^y_T x\in O$.*

Proof. Denote x_c as the attractor for state y. Since O is open, there is an $\varepsilon>0$ such that $\{x':\|x'-x_c\|\leq\varepsilon\}\subset O$. Pick a T such that $\|x-x_c\|C_\gamma\exp(-\gamma T)\leq\varepsilon$, then it suffices to see that

$$\|\Psi^y_T x-x_c\|=\|\Psi^y_T x-\Psi^y_T x_c\|\leq\|x-x_c\|\int_0^1\|D_x\Psi^y_T(x_c+s(x-x_c))\|ds\leq\varepsilon.$$

□

Lemma 6.6 (Burst of decay). *If $Q_i(z)<0$ for some $z\in E$, then fix any $\varepsilon>0$, we can find an $\xi>0$ such that the following two jump sequences are accessible up to time ε:*

$$\begin{array}{cccccc} \mathbf{t}: & 0 & \xi & \dots & \eta_i\xi & (\eta_i+1)\xi & (\eta_i+2)\xi\\ \eta^1: & \eta & \eta-e_i & \dots & \eta-\eta_ie_i & \textit{no jump} & \textit{no jump}\\ \eta^2: & \eta & \eta-e_i & \dots & \eta-\eta_ie_i & \eta-\eta_ie_i+e_i & \textit{no jump}\end{array}$$

while the following hold:

$$(\eta_i+2)\xi\leq\varepsilon,\quad\|\Psi(x,\mathbf{t},\eta^k,s)-x\|\leq\varepsilon,\quad\forall s\leq(\eta_i+2)\xi,k=1,2.$$

Also note this claim also holds when $\eta_i=0$.

Proof. Based the construction of Lemma 6.4, it suffices for us to notice that $\lambda(x,\eta-ke_i,\eta-(k+1)e_i)=c_d(\eta_i-k,Q_i)>0$ for $k<\eta_i$, $\lambda(x,\eta-\eta_ie_i,\eta-(\eta_i-1)e_i)=c_a(0,Q_i)>0$, and the pseudo jump rate $\lambda(x,y,y)=1$. □

Lemma 6.7. *The attractor of the stochastic skeleton model* (6.1) *with given A_i, or equivalently $D_i=S^\theta_i-\bar{H}A_i$ is given by the following:*

$$\begin{aligned} K_i&=\sum_{k=0}^{N-1}\frac{N(1+N\bar{d})^{N-k-1}}{2((1+N\bar{d})^N-1)}D_{i-k},\\ R_i&=\sum_{k=0}^{N-1}\frac{N(1+3N\bar{d})^{N-k-1}}{((1+3N\bar{d})^N-1)}D_{i+k},\quad Z_i=\frac{D_i(1-\bar{Q})}{\bar{d}}.\end{aligned}\tag{6.4}$$

Proof. Since the attractor for each combination of D_i is unique, it suffices for us to find one equilibrium point of ODE system (2.7). As the Z_i parts are independent of others, it is simple to obtain the result. For the K_i part, we look for solutions of following form due to the shift invariant nature of $I=\mathbb{Z}/N\mathbb{Z}$ and (2.7):

$$K_i=\sum_{k=0}^{N-1}a_kD_{i+k}\Rightarrow K_{i-1}=\sum_{k=0}^{N-1}a_{k+1}D_{i+k}.$$

Plug this solution into (2.7), and equating the coefficient of each D_i, we have

$$\begin{cases} (\frac{1}{N}+\bar{d})a_k = \frac{1}{N}a_{k+1}, & k \neq 0 \\ (\frac{1}{N}+\bar{d})a_0 = \frac{1}{2}+\frac{1}{N}a_1 \end{cases}$$

Hence we can obtain that $a_k = (1+N\bar{d})^{k-1}a_1$ and $a_0 = (1+N\bar{d})^{N-1}a_1$, which eventually leads to (6.4) with a similar formula for R_i. □

Now we are finally at the position to prove Lemma 6.3.

Proof of Lemma 6.3. In the following, we will use symbol Z to denote the pair (U,η). We will denote the accessible set from a point $z=(U,\eta)$ as

$$\mathscr{A}^z = \{(\Psi(z,\mathbf{y},\mathbf{t},t),y_n) : p^{z,t}_{n,\mathbf{t},\mathbf{y}} > 0, \forall n,\mathbf{t},\mathbf{y},t\}.$$

Recall the product law from Lemma 4.7:

$$p^{z,t+t'}_{n+n',\mathbf{y}+\mathbf{y}',\mathbf{t}+\mathbf{t}'} = p^{z,t}_{n,\mathbf{y},\mathbf{t}}p^{z_t,t'}_{n',\mathbf{y}',\mathbf{t}'}.$$

As a consequence, if $z' \in \mathscr{A}^z$, then $\mathscr{A}^{z'} \subset \mathscr{A}^z$.

According to the illustration in Section 6.3, we split space E into the following four subsets:

$$B_1 = \{z : Q_i(z) < 0, \quad \forall i\}, \quad B_2 = \{z : \exists i \in I, s.t. \quad Q_i(z) > 0\},$$

$$B_3 = \{z : Q_i(z) = 0, \quad \forall i\}, \quad B_4 = \{z : Q_i(z) \leq 0\}/(B_1 \cup B_3).$$

Since $z' \in \mathscr{A}^z$ implies $\mathscr{A}^{z'} \subset \mathscr{A}^z$, it suffices to show the following claims:

(1) $\vec{1} \in \mathscr{A}^z, \forall z \in B_1$;
(2) $B_1 \cap \mathscr{A}^z \neq \emptyset, \forall z \in B_2$;
(3) $(B_1 \cup B_2) \cap \mathscr{A}^z \neq \emptyset, \forall z \in B_3$;
(4) $(B_1 \cup B_2 \cup B_3) \cap \mathscr{A}^z \neq \emptyset, \forall z \in B_4$.

In other words, we will show that starting from states in B_m, it is possible to reach states in some B_{m-i}, while from B_1 it is possible to reach $\vec{1}$.

Step (1). For $z \in B_1$, consider the following burst sequence from $z \in B_1$ to $\vec{1}$, which is applying the second construction in Lemma 6.6 sequentially at each $i \in I$:

$$\eta,\ldots,\eta-\eta_1 e_1, \eta-(\eta_1-1)e_1, \eta-(\eta_1-1)e_1-e_2,\ldots,$$
$$\ldots,\eta-(\eta_1-1)e_1-(\eta_2-1)e_2,\ldots,\vec{1}.$$

Then by Lemma 6.4, $\vec{1} \in \mathscr{A}^z$.

Step (2). For $z \in B_2$, let i be one of the indices that $Q_i(z) > 0$. Let M be an integer large enough such that if we let

$$D_i = S^\theta_i - M\bar{H}\Delta A, \quad D_j = S^\theta_i - \bar{H}A_j(z), j \neq i$$

then $Q_j = Z_j + \bar{Q}(K_j + R_j)$ with Z_j, K_j, R_j being the attractor given by (6.4) are negative for all $j \in I$. As K_j, R_j depends positively and linearly over D_i in the formulation of (6.4), this M exists. Then as

$$\lambda(U, \eta + ke_i, \eta + (k+1)e_i) = c_a(\eta + ke_i, Q_i) > 0, \quad k \in \mathbb{N},$$

by Lemma 6.4, after a burst sequence at time t_0, $z_{t_0} = (U_{t_0}, \eta_{t_0}) \in \mathscr{A}^z$ with

$$\eta_{t_0} = \eta + Me_i, \quad \|U_{t_0} - U\| \leq 1.$$

Then the attractor of η_{t_0}, by Lemma 6.7 and the choice of K, is in the open set B_1. By Lemma 6.5 there exists an T such that $z_{t_0+T} = (\Psi_T^{y_{t_0}}(z_{t_0}), \eta_{t_0})$ is in B_1.

Step (3). For $z \in B_3$, consider a set valued function $J(z) = \{i \in I : Q_i(z) = 0\}$ and denote its cardinality as $|J(z)|$ and its complement in I as $J^c(z)$. As $|J(z)|$ takes only finitely many values, the minimizer

$$z_0 = arg \min_{z' \in \mathscr{A}^z} |J(z')|$$

can be obtained. Since $|J(z)| \leq N-1$, we have $|J(z_0)| \leq N-1$. Notice that $|J(z_0)| = 0$ implies that $z_0 \in B_1 \cup B_2$, so it suffice to reveal a contradiction in the case that $|J(z_0)| \in [1, N-1]$. As $J(z_0)$ then is neither full nor empty, and I is cyclic, we can find an i such that $i-1 \in J^c(z_0)$ and $i \in J(z_0)$. As Q_j are continuous in $\mathscr{H}$, we can pick any $\varepsilon > 0$ small enough such that the following holds:

$$Q_j(U') < 0, \quad \forall \|U' - U_0\| \leq 2\varepsilon, \quad j \in J^c(z). \tag{6.5}$$

Following Lemma 6.6 we can burst z_0 through either one of the following sequences with a proper ξ

t	0	ξ	$\dots$	$\eta_{i-1}\xi$	$(\eta_{i-1}+1)\xi$	$(\eta_{i-1}+2)\xi$
η^1	η	$\eta - e_{i-1}$	$\dots$	$\eta - \eta_{i-1}e_{i-1}$	no jump	no jump
η^2	η	$\eta - e_{i-1}$	$\dots$	$\eta - \eta_{i-1}e_{i-1}$	$\eta - \eta_{i-1}e_{i-1} + e_{i-1}$	no jump

while the generated **U** part satisfies:

$$\|U_s^1 - U_0\|, \|U_s^2 - U_0\| \leq \varepsilon, \quad s \leq (\eta_{i-1}+2)\xi \leq \varepsilon.$$

Then by (6.5), $J(z_s^1), J(z_s^2) \subseteq J(z_0)$ for $s \leq (\eta_i + 2)\xi$. Since z_0 is a minimizer among $\mathscr{A}^z$, $J(z_s^1) = J(z_s^2) = J(z_0)$, so

$$Q_i(z_s^1) = 0, \quad Q_i(z_s^2) = 0, \quad s \leq (\eta_{i-1}+2)\xi.$$

Then for any $s \in ((\eta_i+1)\xi, (\eta_i+2)\xi)$, the first time differential at s gives us

$$0 = \dot{Q}_i(U_s^k) = [-D_x^+ K_i + D_x^- R_i/3 + D_i(1 - \bar{Q}/6) - \bar{d}Q_i](z_s^k), \quad k = 1, 2.$$

Take time differential again and use $Q_i \equiv 0$ we obtain:

$$\begin{aligned} 0 = \ddot{Q}_i(z_s^k) &= [-D_x^+ D_x^+ K_i - \bar{d}D_x^+ K_i + D_x^- D_x^- R_i/3 - \bar{d}D_x^- R_i](z_s^k) \\ &\quad + [D_x^+ D_i/2 + D_x^- D_i/3](z_s^k) \qquad k = 1, 2. \end{aligned} \tag{6.6}$$

Hence the second line is the opposite of the first line. As the first line of (6.6) is a linear combination of components of U_s^k, so there is a constant M such that the difference of the second line of (6.6) can be bounded as follow:

$$\left|[D_x^+ D_i/2 + D_x^- D_i/3](z_s^1) - [D_x^+ D_i/2 + D_x^- D_i/3](z_s^2)\right| \le M\|U_s^1 - U_s^2\| \le 2M\varepsilon.$$

However, since η_s^1 and η_s^2 differ only at η_{i-1}, with

$$\eta_{i-1}(\eta_s^1) = 0, \eta_{i-1}(\eta_s^2) = 1,$$

so

$$\left|[D_x^+ D_i/2 + D_x^- D_i/3](z_s^1) - [D_x^+ D_i/2 + D_x^- D_i/3](z_s^2)\right|$$
$$= \frac{1}{2N}|A_{i-1}(\eta_s^1) - A_{i-1}(\eta_s^2)| = \frac{\Delta A}{2N}.$$

so if we let ε in addition be less than $\Delta A/4MN$ in the beginning, there will be a contradiction.

Step (4). For $z = (u, \eta) \in B_4$, since the attractor for its η part u_η satisfies $\psi(u_\eta, \eta) = 0$, so by Lemma 2.2 $u_\eta \in B_4^c$. Since B_4^c is an open set, by Lemma 6.5, $B_4^c \cap \mathscr{A}^z \neq \emptyset$. □

7 Concluding Discussion

Stochastic lattice models are prominent ways to capture highly intermittent unresolved features in climate science and material science [28, 30, 14, 11, 22, 24]. Mathematically, they consist of an ODE system X_t and a Markov jump process Y_t, while the evolution of the two depend on each other. Such models are special piecewise deterministic Markov processes (PDMP) [10, 20], while the transition rates of Y_t are sometimes unbounded or contain degeneracy. In order to understand the asymptotic behavior of these models, we develop a general framework, Theorem 3.10, to verify geometric ergodicity under a proper Wasserstein distance. The conditions it requires, heuristically speaking, are: 1) there is a Lyapunov function that controls the transition rates of η_t and their Frechét derivatives; 2) the differential flow of $\mathbf{U}_t$ is piecewise contracting; 3) the process η_t is not reducible. The proof relies on a perturbation analysis of the probability density and applies the asymptotic coupling framework in [16, 18]. Since the techniques used here rely more on analysis rather than the concrete coupling construction, the proofs appear to be more straight forward and easier to generalize than previous treatments of PDMP.

In order to demonstrate the applications of our results, Theorem 3.10 is applied to two stochastic lattice models from the existing climate science literature. The application to the simplest tropical climate model [15, 32] is rather straight forward, as its energy is path-wise dissipative, hence the dynamics is contained in a compact invariant set. The application to the skeleton model for MJO [40, 39, 33], on the other hand, is much more non trivial, since its energy is dissipative only on average and does not regularize the transition rates directly, moreover the transition rates can be degenerate. These difficulties are resolved by considering a higher moment of the energy function, and running a hypo-elliptic type of verification through the vector fields. This interesting application demonstrates the power and versatility of our framework.

Despite the fact that Theorem 3.10 allows unbounded degenerate transition rates, the requirement that the differential flow is piecewise contracting constrains us from more general applications. A more general setting will be assuming the differential flow contracts and expands with a rate that depends on the jump process Y_t, while on average the dynamics is contracting [1, 4, 7]. One potential extension of this paper will be showing geometric ergodicity for unbounded transition rates with the differential flow being contracting only on average. Yet this cannot be carried out simply by upgrading the bounds in our proofs. The reason can be illustrated through a comparison between the Wasserstein distance we used here with the one of [7]. Through the asymptotic coupling framework of [18], the contracting distance we constructed in (4.5) is roughly of form $\int_x^{x'} V du$, while the distance used in [7] is roughly $\|x-x'\|^q$, where q can be a number less than 1. Neither one of the two dominates the other. Hence, in order to extend our results to contracting on average dynamics, a systematical upgrade is required. Since the major goal of the current paper is to find a general framework for stochastic lattice models, this interesting extension will be carried out in another paper of the authors [41].

Another promising method to show geometric ergodicity is through a hypoelliptic argument, as shown in [3]. Although it seems not difficult to generalize the corresponding results to processes with unbounded transition rates, one major constraint is the verification of the Hörmander condition. The vector fields for stochastic lattice models, like the ones introduced in Section 2, are usually of dimension 50 or more, while all dimensions are correlated through the neighboring interaction; a direct verification of the Hörmander condition by hand seems relatively impossible. A theorem that can simplify this verification will be very interesting. Another alternative would be developing algorithms for such purpose.

Acknowlegdements

This research is supported by the MURI award grant N-000-1412-10912, where A.J.M. is the principal investigator, while X.T.T. is supported as a postdoctoral fellows.

Appendix A: Miscellaneous Results

Lemma A.1. *For two probability measures μ, ν on a polish space, assume that $\mu(X \in A) \geq p$ and $\nu(Y \in B) \geq p$, then there exists a coupling Γ of μ and ν such that $\Gamma(X \in A, Y \in B) \geq p$.*

Proof. Consider adding an independent Bernoulli random variable W into the the probability space of μ, such that $\mu(W = 1) = p/\mu(X \in A)$. Then if we let $X' = (X, W)$ and $A' = \{(x, 1) : x \in A\}$, then $\mu(X' \in A') = p$. So without loss of generality, we can assume $\mu(X \in A) = \nu(Y \in B) = p$. Define Γ through the following:

$$\begin{aligned}\Gamma(X \in C, Y \in D) = \mu(X \in C | X \in A)\nu(Y \in D | Y \in B)p \\ + \mu(X \in C | X \in A^c)\nu(Y \in D | Y \in B^c)(1 - p).\end{aligned}$$

Let C or D be the whole space, one can easily verifies that Γ is a coupling of μ and ν; let $C = A$ and $D = B$, it is clear that Γ satisfies our requirement. □

Appendix B: Dissipation of Energy

Lemma B.1. *Assuming relation* (2.5)*, the energy density of the simplest tropical climate model given by* (5.3) *follows a path-wise dissipative principle:*

$$\frac{d\varepsilon_i}{dt} \leq -K_i D_x K_i + R_i D_x^- R_i - \gamma\varepsilon_i + k_v(\theta_{s,i}, \theta_{eq,i}, q_{s,i})$$

with a proper $\gamma, k_v > 0$.

Proof. The time derivative of ε_i is:

$$\begin{aligned}\frac{d\varepsilon_i}{dt} &= -K_i D_x^+ K_i + R_i D_x^- R_i - \frac{(\bar{d} + d_\theta + d_{sh})}{2}(K_i^2 + R_i^2) - \frac{2d_q}{(\alpha + \bar{Q})(1 - \bar{Q})} Z_i^2 \\ &\quad - (d_\theta + d_{sh} - \bar{d})K_i R_i + (d_\theta \theta_{eq,i} + d_{sh}\theta_{s,i})(K_i + R_i) - \frac{2(q_i - \alpha\theta_i)P_i}{(\alpha + \bar{Q})(1 - \bar{Q})} \\ &\quad + \frac{2(d_q q_{s,i} + \bar{Q} d_\theta \theta_{eq,i} + \bar{Q} d_{sh} q_{s,i})Z_i}{(\alpha + \bar{Q})(1 - \bar{Q})} + \frac{(d_\theta + d_{sh} - d_q)\bar{Q}(K_i + R_i)Z_i}{(\alpha + \bar{Q})(1 - \bar{Q})}.\end{aligned}$$

First, notice that

$$\begin{aligned}-\frac{(\bar{d} + d_\theta + d_{sh})}{2}(K_i^2 + R_i^2) - (d_\theta + d_{sh} - \bar{d})K_i R_i \\ = -\frac{(d_\theta + d_{sh})}{2}(K_i + R_i)^2 - \frac{\bar{d}}{2}(K_i - R_i)^2.\end{aligned}$$

With (2.5):

$$(1 - \bar{Q})(\alpha + \bar{Q})d_q(d_\theta + d_{sh}) \geq (d_\theta + d_{sh} - d_q)^2 \bar{Q}^2,$$

the linear and cross terms are bounded by:

$$\frac{(d_\theta+d_{sh}-d_q)\bar{Q}(K_i+R_i)Z_i}{(\alpha+\bar{Q})(1-\bar{Q})}\leq\frac{d_\theta+d_{sh}}{4}(K_i+R_i)^2+\frac{d_q}{(\alpha+\bar{Q})(1-\bar{Q})}Z_i^2,$$

$$(d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i})(K_i+R_i)\leq\frac{d_\theta+d_{sh}}{8}(K_i+R_i)^2+\frac{2(d_\theta\theta_{eq,i}+d_{sh}\theta_{s,i})^2}{d_\theta+d_{sh}},$$

$$\frac{2(d_qq_{s,i}+\bar{Q}d_\theta\theta_{eq,i}+\bar{Q}d_{sh}\theta_{s,i})Z_i}{(\alpha+\bar{Q})(1-\bar{Q})}\leq\frac{2(d_qq_{s,i}+\bar{Q}d_\theta\theta_{eq,i}+\bar{Q}d_{sh}\theta_{s,i})^2}{d_q(\alpha+\bar{Q})(1-\bar{Q})}+\frac{d_qZ_i^2}{2(\alpha+\bar{Q})(1-\bar{Q})}$$

Aslo, notice that $P_i(q_i-\alpha\theta_i)\geq 0$ since $P_i=(1-\sigma_I)\tau_c^{-1}(q_i-\alpha\theta_i-\hat{q})$, so in combine

$$\begin{aligned}\frac{d\varepsilon_i}{dt}&\leq -K_iD_x^+K_i+R_iD_x^-R_i-\frac{\bar{d}}{2}(K_i-R_i)^2-\frac{d_\theta+d_{sh}}{8}(K_i+R_i)^2\\&\quad-\frac{d_q}{2(\alpha+\bar{Q})(1-\bar{Q})}Z_i^2+\bar{d}^{-1}(d_\theta\theta_{eq,i}^2+d_{sh}\theta_{s,i}^2)\\&\quad+\frac{4(d_qq_{s,i}+\bar{Q}d_\theta\theta_{eq,i}+\bar{Q}d_{sh}\theta_{s,i})^2}{d_q(\alpha+\bar{Q})(1-\bar{Q})}\\&\leq -K_iD_x^+K_i+R_iD_x^-R_i-\gamma\varepsilon_i+k_v(\theta_{s,i},\theta_{eq,i},\theta_{eq,i}).\end{aligned}$$

where $\gamma=\min\{\bar{d},(d_\theta+d_{sh})/4,\frac{d_q}{2}\}$, and we used

$$-\frac{\bar{d}}{2}(K_i-R_i)^2-\frac{d_\theta+d_{sh}}{8}(K_i+R_i)^2\leq-\frac{\gamma}{2}[(K_i-R_i)^2+(K_i+R_i)^2]=-\gamma(K_i^2+R_i^2).$$

□

Lemma B.2. *Assuming relation* (2.5)*, following the derivative flow* (5.5)*, the* ε_i^h *defined by* (5.6) *has the following dissipative behavior with some proper* $\gamma>0$*:*

$$\frac{d\varepsilon_i^h}{dt}\leq -K_i^hD_x^+K_i^h+R_i^hD_x^-R_i^h-\gamma\varepsilon_i^h.$$

Proof. following the derivation of Lemma B.1, we find

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}((K_i^h)^2+(R_i^h)^2)=&-K_i^hD_x^+K_i^h+R_i^hD_x^-R_i^h-\frac{d_\theta+d_{sh}}{2}(K_i^h-R_i^h)^2\\&-\frac{\bar{d}}{2}(K_i^h+R_i^h)^2-P_i^h(K_i^h+R_i^h),\end{aligned}$$

$$\frac{d}{dt}\frac{(Z_i^h)^2}{(1-\bar{Q})(\alpha+\bar{Q})}=-\frac{2P_i^hZ_i^h}{\alpha+\bar{Q}}-\frac{2d_q(Q_i^h)^2}{(1-\bar{Q})(\alpha+\bar{Q})}+\frac{\bar{Q}Q_i^h(d_\theta+d_{sh}-d_q)(K_i^h+R_i^h)}{(1-\bar{Q})(\alpha+\bar{Q})},$$

With relation (2.5), i.e.

$$(1-\bar{Q})(\alpha+\bar{Q})d_q(d_\theta+d_{sh})\geq(d_\theta+d_{sh}-d_q)^2\bar{Q}^2,$$

Young's inequality can be applied to the cross term,

$$\frac{\bar{Q}Q_i^h(d_\theta+d_{sh}-d_q)(K_i^h+R_i^h)}{(1-\bar{Q})(\alpha+\bar{Q})} \le \frac{d_q(Q_i^h)^2}{(1-\bar{Q})(\alpha+\bar{Q})} + \frac{d_\theta+d_{sh}}{4}(K_i^h+R_i^h)^2.$$

Also notice that when $q_i-\alpha\theta_i-\hat{q}>0$ or $q_i-\alpha\theta_i-\hat{q}=0$ and $q_i^h-\alpha\theta_i^h>0$, then

$$(K_i^h+R_i^h+\frac{2Z_i^h}{\alpha+\bar{Q}})P_i^h = \frac{1}{2}(1-\sigma_I)\tau_c^{-1}(K_i^h+R_i^h+\frac{2Z_i^h}{\alpha+\bar{Q}})^2 \ge 0;$$

else $(K_i^h+R_i^h+\frac{2Z_i^h}{\alpha+\bar{Q}})P_i^h=0$, in other words it is not negative. Hence we have

$$\begin{aligned}\frac{d\varepsilon_i^h}{dt} &\le -K_i^hD_x^+K_i^h+R_i^hD_x^-R_i^h-\frac{d_\theta+d_{sh}}{4}(K_i^h+R_i^h)^2\\ &\quad -\frac{\bar{d}}{2}(K_i^h-R_i^h)^2-\frac{d_q(Q_i^h)^2}{2(1-\bar{Q})(\alpha+\bar{Q})}\\ &\le -K_i^hD_x^+K_i^h+R_i^hD_x^-R_i^h-\gamma\varepsilon_i^h,\end{aligned}$$

with $\gamma=\min\{\bar{d},(d_\theta+d_{sh})/2,\frac{1}{2}d_q\}$. □

Lemma B.3. *For the sytem* (6.1)*, denote*

$$\varepsilon_i=\frac{1}{2}\left[2K_i^2+3R_i^2+\frac{(Z_i+1)^2}{(1-\bar{Q})\bar{Q}}\right]+\frac{\bar{H}A_i}{\Gamma\bar{Q}},$$

then the following hold with a proper $\gamma,k_v>0$:

$$\mathscr{L}\varepsilon_i \le -2K_iD_x^+K_i+R_iD_x^-R_i-\gamma\varepsilon_i+k_v(S_i^\theta).$$

Proof. Denote $D_i=S_i^\theta-\bar{H}A_i$. Notice that

$$\begin{aligned}\frac{d}{dt}K_i^2 &= -2K_iD_x^+K_i+K_iD_i-2\bar{d}K_i^2;\\ \frac{d}{dt}R_i^2 &= \frac{2}{3}R_{i+1}D_xR_i+\frac{2}{3}R_iD_i-2\bar{d}R_i^2.\end{aligned}$$

Morover,

$$\frac{d}{dt}(Z_i+1)^2=2(D_i(1-\bar{Q})+\bar{d})(Z_i+1)-2\bar{d}(Z_i+1)^2.$$

As

$$\mathscr{L}A_i=\Gamma Q_iA_i+\Delta A1_{A_i=0}=\frac{\Gamma}{\bar{H}}((Z_i+1)+\bar{Q}(K_i+R_i)-1)(S_i^\theta-D_i)+\Delta A1_{A_i=0}$$

then in combine

$$\begin{aligned}\mathscr{L}\varepsilon_i &= -2K_iD_x^+K_i+R_iD_x^-R_i-2\bar{d}K_i^2-3\bar{d}R_i^2-\frac{\bar{d}(Z_i+1)^2}{(1-\bar{Q})\bar{Q}}-\frac{\bar{H}A_i}{\bar{Q}}\\ &\quad +\frac{S_i^\theta}{\bar{Q}}+\bar{d}(Z_i+1)+\frac{S_i^\theta}{\bar{Q}}((Z_i+1)+\bar{Q}(K_i+R_i)-1)-\frac{\bar{H}\Delta A}{\Gamma\bar{Q}}1_{A_i=0}\end{aligned}$$

Applying Young's inequality:

$$\left(\bar d+\frac{S_i^\theta}{\bar Q}\right)(Z_i+1)\le \frac{\bar d(Z_i+1)^2}{4(1-\bar Q)\bar Q}+(1-\bar Q)\bar Q d^{-1}\left(\bar d+\frac{S_i^\theta}{\bar Q}\right)^2,$$

$$\begin{aligned}S_i^\theta\left(\frac{1+Z_i}{\bar Q}+(K_i+R_i)\right)\le&\left(\frac{\bar d(1+Z_i)^2}{4(1-\bar Q)\bar Q}+\frac{(S_i^\theta)^2(1-\bar Q)}{\bar d\bar Q}\right)\\&+\left(\bar d K_i^2+\frac{(S_i^\theta)^2}{4\bar d}\right)+\left(2\bar d R_i^2+\frac{(S_i^\theta)^2}{8\bar d}\right).\end{aligned}$$

With these in hand we have

$$\begin{aligned}\partial_t\varepsilon_i\le&-2K_iD_x^+K_i+R_iD_x^-R_i-\bar dK_i^2-\bar dR_i^2-\frac{\bar d(Z_i+1)^2}{2(1-\bar Q)\bar Q}-\frac{\bar HA_i}{\bar Q}\\&+(1-\bar Q)\bar Q d_z^{-1}\left(\bar d+\frac{S_i^\theta}{\bar Q}\right)^2+(S_i^\theta)^2\left[\frac{(1-\bar Q)}{\bar d\bar Q}+\frac{1}{4\bar d}+\frac{1}{8\bar d}\right]+\frac{S_i^\theta}{\bar Q}\\\le&-2K_iD_x^+K_i+R_iD_x^-R_i-\gamma\varepsilon+k_v(S_i^\theta)\end{aligned}$$

with $\gamma=\frac{2}{3}\bar d\wedge\Gamma$. $\square$

Bibliography

[1] Bardet, J. B.; Guérin, H.; Malrieu, F. Long time behavior of diffusions with markov switching. *ALEA Lat. Am. J. Probab. Math. Stat.* **7** (2010), 151– 170.

[2] Benaïm, M.; Hurth, T. Invariant densities for dynamical systems with randomswitching. *Nonlinearity* **25** (2012), no. 10, 2937–2952.

[3] Benaïm, M.; Le Borgne, S.; Malrieu, F.; Zitt, P. A. Qualitative properties of certain piewise deterministic markov process. 2012. ArXiv:1204.4143.

[4] Benaïm, M.; Le Borgne, S.; Malrieu, F.; Zitt, P. A. Quantitative ergodicity for some switched dynamical systems. *Electron. Commun. Probab.* **17** (2012), no. 56, 1–14.

[5] Betts, A. K. A new convective adjustment scheme, part i: Observational and theoretical basis,. *Quart. J. Roy. Meteor. Soc.* **112** (1986), 677–692.

[6] Betts, A. K.; Miller, M. J. A new convective adjustment scheme, part ii: Single column tests using gate wave, bomex, and arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.* **112** (1986), 693–709.

[7] Cloez, B.; Hairer, M. Exponential ergodicity for markov processes with random switching. 2014. Accepted by Bernoulli, arXiv:1303.6999.

[8] Da Prato, G.; Zabczyk, J. *Ergodicity for infinite dimensional systems*, London mathematical society lecture note series, Cambridge University Press, 2006.

[9] Davis, M. H. A. Piecewise-deterministic markov processes: A general class of non-diffusion stochastic models. *J. R. Statist. Soc. B.* **46** (1984), no. 3, 353–388.

[10] Davis, M. H. A. *Markov Models and Optimization*, Monographs on Statisitics and Applied Probability, Chapman & Hall, 1993.

[11] Deng, Q.; Khouider, B.; Majda, A. J. The mjo in a coarse-resolution gcm with a stochastic multicloud parameterization. 2014. Submitted to J. Atoms. Sci.

[12] Dudley, R. M. *Real Analysis and Probability*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2004.

[13] Frenkel, Y.; Majda, A. J.; Khouider, B. Using the stochastic multicloud model to improve tropical convective parameterization: A paradigm example. *J. Atmos. Sci.* **69** (2012), no. 3, 1080–1105.

[14] Frenkel, Y.; Majda, A. J.; Khouider, B. Stochastic and deterministic multicloud parameterizations for tropical convection. *Climate Dynamics* **41** (2013), no. 5-6, 1527–1551.

[15] Frierson, D.; Majda, A. J.; Pauluis, O. M. Large scale dynamics of precipitation fronts in the tropical atmosphere: A novel relaxation limit. *Communiations in Mathematical Sciences* **2** (2004), no. 4, 591–626.

[16] Hairer, M.; Mattingly, J. C. Ergodicity of the 2d navier-stokes equations with degnerate stochastic forcing. *Annals of Mathematics* **164** (2006), 993–1032.

[17] Hairer, M.; Mattingly, J. C. A theory of hypoellipticity and unique ergodicity for semilinear stochastic pdes. *Electron. Commun. Probab.* **16** (2011), no. 23, 658–738.

[18] Hairer, M.; Mattingly, J. C.; Scheutzow, M. Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations. *Probab. Theory Related Fields* **149** (2011), no. 1-2, 223–259. Available at: `http://dx.doi.org/10.1007/s00440-009-0250-6`

[19] Hirsch, C. *Numerical Computation of Internal and External Flows: The Fundamentals of Computational Fluid Dynamics, Second Edition*, The fundamentals of computational fluid dyanmics, Butterworth-Heinemann, 2007.

[20] Jacobson, M. *Point Process Theory and Applications Marked POint and Piecewise Deterministic Processes*, Probability and its Applications, Birkhauser, 2006.

[21] Katsoulakis, M. A.; Majda, A. J.; Sopasakis, A. Multiscale couplings in prototype hybrid deterministic/stochastic systems: Part i, deterministic closures. *Comm. Math. Sci.* **2** (2004), no. 2, 255–294.

[22] Katsoulakis, M. A.; Majda, A. J.; Sopasakis, A. Multiscale couplings in prototype hybrid deterministic/stochastic systems: Part ii, stochastic closures. *Comm. Math. Sci.* **3** (2005), no. 3, 453–478.

[23] Katsoulakis, M. A.; Majda, A. J.; Sopasakis, A. Intermittency, metastability and coarse graining for coupled deterministic-stochastic lattice systems. *Nonlinearity* **19** (2006), 1021–1323.

[24] Katsoulakis, M. A.; Majda, A. J.; Sopasakis, A. Hybrid deterministic stochastic systems with microscopic look-ahead dynamics. *Comm. Math. Sci.* **8** (2010), no. 2, 409–437.

[25] Katsoulakis, M. A.; Majda, A. J.; Vlachos, G. D. Coarse-grained stochastic processes and monte-carlo simulations in lattice systems. *J. Comput. Phys* **186** (2003), 250–278.

[26] Khouider, B. A coarse grained stochastic multi-type particle interacting model for tropical convection: Nearest neighbour interactions. *Comm. Math. Sci.* **12** (2014), no. 8, 1379–1407.

[27] Khouider, B.; Biello, J. A.; Majda, A. J. A stochastic multicloud model for tropical convection. *Comm. Math. Sci.* **8** (2010), no. 1, 187–216.

[28] Khouider, B.; Majda, A. J.; Katsoulakis, M. A. Coarse-grained stochastic models for tropical convection and climate. *Proc. Natl. Acad. Sci.* **100** (2003), no. 21, 11 941–11 946.

[29] Khouider, B.; Majda, A. J.; Stechmann, S. Climate science in the tropics: Waves, vortices, and pdes. *Nonlinearity* **26** (2013), no. R1-R68.

[30] Majda, A. J.; Franzke, C.; Khouider, B. An applied mathematics perspective on stochastic modelling for climate. *Phil. Trans. Roy. Soc.* **336** (2008), no. 1875, 2427–2453.

[31] Majda, A. J.; Khouider, B. Stochastic and mesoscopic models for tropical convection. *Proc. Natl Acad. Sci.* **99** (2002), 1123–1128.

[32] Majda, A. J.; Souganidis, P. The existence and uniqueness of weak solutions for precipitation fronts: A novel hyperbolic free boundary problem in several space variables. *Comm. Pure Appl. Math* **63** (2010), no. 10, 1351–1361.

[33] Majda, A. J.; Stechmann, S. The skeleton of tropical intraseasonal oscillations. *Proc. Natl. Acad. Sci.* **106** (2009), no. 21, 8417–8422.

[34] Majda, A. J.; Stechmann, S. Nonlinear dynamics and regional variations in the mjo skeleton. *J. Atmos. Sci.* **68** (2011), no. 12, 3053–3071.

[35] Mao, X.; Yuan, C. Asymptotic stability in distribution of stochastic differential equations with Markovian switching. *Stochastic processes and their applications* **103** (2003), 277–291.

[36] Mattingly, J. C.; Stuart, A. M. Geometric ergodicity of some hypo-elliptic diffustion for particle motions. *Markov Process. Related Fields* **8** (2002), no. 2, 199–214.

[37] Meyn, S.; Tweedie, R. *Markov chains and stochastic stability*, Springer-Verlag, 1993.

[38] Neelin, J.; Zeng, N. A quasi-equilibrium tropical circulation model: Formulation. *J. Atmos. Sci.* **57** (2000), 1741–1766.

[39] Thual, S.; Majda, A. J.; Stechmann, S. Assymetrical intraseasonal events in the stochastic skeleton mjo model with seasonal cycle. 2014. Accepted by Climate Dynamics.

[40] Thual, S.; Majda, A. J.; Stechmann, S. A stochastic skeleton model for mjo. *J. Atmos. Sci.* **71** (2014), 697–715.

[41] Tong, X. T.; Majda, A. J. Geometric ergodicity for diffusions with random regimes switching. In preparation.

Received Month 200X.