A Necessary and Sufficient Condition for Global Stability of Monotone Markov Processes

Takashi Kamihigashi[a], John Stachurski[b,*]

[a]*Center for Computational Social Science, Kobe University, Kobe, Japan*
[b]*National Graduate Institute for Policy Studies, Tokyo, Japan*

ARTICLE INFO

Keywords:
global stability
monotone Markov process
stochastic dominance
ergodicity
2020 MSC: 60J05, 60J25, 37A30, 47H10, 60E15

ABSTRACT

We characterize global stability of monotone Markov processes, showing that any such process has a unique, globally stable stationary distribution if and only if it is asymptotically contractive and has a tight trajectory. The characterization follows from a new fixed point theorem for order-preserving maps on preordered metric spaces, and covers compact and noncompact state spaces and both discrete and continuous time. We illustrate our results with applications to Bayesian learning and income dynamics.

1. Introduction

Stochastically monotone Markov processes—processes for which higher current states induce stochastically higher future states—are a classical object of study in probability theory [1–4]. A central question is when such a process admits a unique and globally stable stationary distribution. Monotonicity is valuable here because it opens the door to order-based arguments, often involving coupling [5–9]. Early results established global stability under relatively strong assumptions that induce a uniform rate of contraction to the stationary distribution [10, 11]. A more recent literature weakens the mixing requirements, replaces compactness with tightness-type conditions, and accommodates non-irreducible and partially monotone chains (see, e.g., [12–20]). The conditions obtained in this way are sufficient but not necessary.

This paper provides simple necessary and sufficient conditions for global stability of monotone Markov processes on compact or noncompact state spaces. The conditions rest on a new fixed point theorem for order-preserving maps on preordered metric spaces. The theorem states that, on a complete preordered metric space whose metric satisfies a diagonal property, an order-preserving map has a unique fixed point—necessarily globally stable—if and only if it is asymptotically contractive and generates at least one order-bounded trajectory. We also provide an analogous result for asymptotically contractive and order-preserving semigroups that can be used to handle continuous time.

In the Markov setting, asymptotic contractivity holds under weak mixing conditions that generalize conditions of [10, 11]. This asymptotic-contractivity route to stability parallels the asymptotic-coupling approach to ergodicity without irreducibility, developed for general Markov processes [21] and for order-preserving semigroups [22]. Moreover, we show that when distributions are ordered by stochastic dominance, existence of an order-bounded trajectory is equivalent to existence of a tight trajectory—a weak form of boundedness in probability, which is a standard stability condition in the theory of Markov processes (see, e.g., [23–25]), and one considerably weaker than compactness of the state space. The result is a sharp characterization: a monotone Markov semigroup is globally stable if and only if it is asymptotically contractive and has a tight trajectory.

These results are established in an abstract semigroup setting that treats continuous and discrete time on an equal footing, extending the classical theory to continuous-time models. The framework also accommodates nonlinear Markov operators, as arise in distribution dynamics with aggregate feedback (see, e.g., [19]).

On a theoretical level, the closest antecedents are [13, 17], which establish global stability from a combination of (a) order-theoretic mixing conditions and (b) side conditions that imply existence of a stationary distribution. Under matching assumptions on the underlying state space, both results are special cases of the abstract stability theorem obtained here (Theorem 2.1), which shows that asymptotic contractivity together with a single order-bounded trajectory is enough for global stability. The present theorem goes further: it admits operators on arbitrary preordered spaces,

*Corresponding author
`tkamihig@rieb.kobe-u.ac.jp` (T. Kamihigashi); `j-stachurski@grips.ac.jp` (J. Stachurski)
ORCID(s):

supplies necessary as well as sufficient conditions, and covers continuous time alongside discrete time. The discrete-time results of [17] are complementary, allowing for weaker conditions on the state space and contributing valuable sufficient conditions for order boundedness and asymptotic contractivity.

Our fixed point theorem belongs to a broad literature on fixed points of order-preserving maps on spaces with metric or topological structure. A seminal contribution is [26], an analogue of Banach's fixed point theorem for order-preserving maps in which the contraction condition need hold only on comparable elements; this line has since been developed in many directions (see, e.g., [27–30]). Our approach differs from this literature in two respects: we use no algebraic structure, and we impose asymptotic contractivity directly rather than one- or multi-step contraction conditions. Although asymptotic contractivity is less primitive than a one-step contraction, it is well suited to the Markov setting, where it follows naturally from coupling arguments: under weak monotone mixing conditions, two trajectories started from different initial conditions are drawn together in the Bhattacharya metric.

Two applications illustrate the scope of the theory. The first is a model of Bayesian learning subject to belief shocks, as arise under countercyclical uncertainty and forecast bias [31, 32]. The second treats income dynamics generated by piecewise deterministic Markov processes—combining deterministic drift with random jumps—which produce Pareto tails in the stationary distribution, complementing earlier work in the same direction [33–36]. We also show that, under the stronger classical assumptions of a compact state space with least and greatest elements, the framework yields a uniform exponential rate of convergence and ergodicity, recovering and sharpening the monotone mixing result of [11].

The remainder of the paper is organized as follows. Section 2 develops the abstract fixed point theory for order-preserving semigroups. Section 3 specializes to probability spaces, introducing the Bhattacharya metric and connecting tightness to order boundedness. Section 4 covers stochastic kernels and transition probability functions. Section 5 establishes verifiable conditions for global stability of monotone Markov processes in discrete and continuous time, and, under stronger assumptions on the state space, uniform exponential convergence and ergodicity. Section 6 develops a stability result for piecewise deterministic Markov processes. Section 7 contains two applications: Bayesian learning with belief shocks and income dynamics with deterministic drift. Section 8 concludes. Proofs are collected in Appendix A.

2. Stability of Order-Preserving Semigroups

This section provides our high-level results on stability of semigroups, while also providing, as a corollary, a new fixed point result for order-preserving maps on preordered metric spaces. The section begins with background on preordered metric spaces and the diagonal property. We then introduce semigroups and the notions of asymptotic contractivity and global stability. At the end we state our main results.

2.1. Background

Let X be a nonempty set. If A is any subset of X, then A^c refers to its complement in X. Given a preorder $\preceq$ on X (reflexive, transitive), each subset I of X having the form

$$I := [a, b] := \{x \in \mathsf{X} \,:\, a \preceq x \preceq b\} \qquad (a, b \in \mathsf{X},\ a \preceq b)$$

is called an *order interval* in X. A set $B \subset \mathsf{X}$ is called *order-bounded* if there exists an order interval I in X with $B \subset I$. A sequence $(x_n) := (x_n)_{n\in\mathbb{N}} \subset \mathsf{X}$ is called *increasing* if $x_n \preceq x_{n+1}$ for all n and *decreasing* if $x_{n+1} \preceq x_n$ for all n. A self-map $T : \mathsf{X} \to \mathsf{X}$ is called *order-preserving* if $x \preceq y$ implies $Tx \preceq Ty$.

A *preordered metric space* is a triple $(\mathsf{X}, d, \preceq)$ where (X, d) is a metric space and $\preceq$ is a preorder on X. For such a space, the preorder $\preceq$ is called *closed* if $(x_i) \subset \mathsf{X}$, $(y_i) \subset \mathsf{X}$, $x_i \preceq y_i$ for all i, $x_i \to x \in \mathsf{X}$ and $y_i \to y \in \mathsf{X}$ implies $x \preceq y$. We say that d has the *diagonal property* with respect to $\preceq$ if, for all a, b in X,

$$d(x, y) \le d(a, b) \quad \text{whenever} \quad x, y \in [a, b]. \tag{2.1}$$

Example 2.1. Let X be a Banach lattice with norm $\|\cdot\|$, partial order $\le$, and absolute value function $|\cdot|$. Fix a, b in X and $x, y \in [a, b]$. Since $-x \le -a$ and $y \le b$, we have $y - x \le b - a$. A similar argument gives $x - y \le b - a$. Hence $|x - y| \le b - a = |b - a|$. Norms on Banach lattices preserve absolute order, so $\|x - y\| \le \|a - b\|$.

Example 2.2. Let A be any set, let X be a set of real-valued functions on A, and let $\preceq$ be the pointwise order on X. If d has the form $d(x, y) = e(|x - y|)$ for some non-decreasing function e mapping into $\mathbb{R}_+ := [0, \infty)$, then d satisfies the diagonal property. Indeed, if we fix $x, y, a, b \in \mathsf{X}$ with $x, y \in [a, b]$, then $a - b \preceq x - y \preceq b - a$ and hence $|x - y| \preceq |b - a|$. Since e is non-decreasing on X, this yields $d(x, y) = e(|x - y|) \le e(|b - a|) = e(|a - b|) = d(a, b)$.

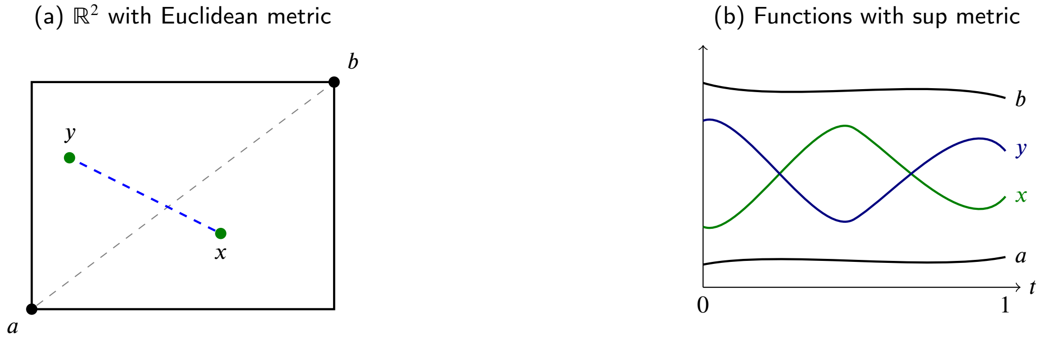

Figure 1: Diagonal property: if $x, y \in [a, b]$, then $d(x, y) \leq d(a, b)$

Examples 2.1 and 2.2 show that the diagonal property also holds in most of the classical spaces and any of their subsets (endowed with the same metric and preorder). Obvious special cases include $\mathsf{X} = \mathbb{R}^n$, where d is Euclidean distance and $\preceq$ is the usual pointwise partial order; X = all bounded, real-valued functions on a given set M, where d is the supremum distance and $\preceq$ is the pointwise partial order; and X is an L_p space, d is the L_p distance, and $\preceq$ is the almost everywhere pointwise partial order (see, e.g., [37]). Figure 1 helps illustrate the diagonal property in the first two cases.

2.2. Semigroups

Let $(\mathsf{X}, d, \preceq)$ be a preordered metric space and let $(T_t) := (T_t)_{t\in\mathbb{T}}$ be a family of self-maps on X, where $\mathbb{T}$ is either $\mathbb{R}_+$ or $\mathbb{Z}_+$. The case $\mathbb{T} = \mathbb{R}_+$ represents continuous time, while $\mathbb{T} = \mathbb{Z}_+$ represents discrete time. The family (T_t) is called a *semigroup* on X when $T_0 = I$ and $T_{s+t} = T_s \circ T_t$ for all $s, t \in \mathbb{T}$. We refer to $\mathbb{T}$ as the *index set*. (Unlike some studies of operator semigroups, no algebraic structure is imposed on the underlying set X, and we do not require that $t \mapsto T_t$ is continuous.)

Given $x \in \mathsf{X}$, the map $t \mapsto T_t x$ from $\mathbb{T} \to \mathsf{X}$ is called the *trajectory* of x under the semigroup (T_t). The trajectory is called *order-bounded* if its range is contained in an order interval; that is, if there exist $a, b \in \mathsf{X}$ such that $a \preceq T_t x \preceq b$ for all $t \in \mathbb{T}$. We say that $x^* \in \mathsf{X}$ is a *stationary point* of (T_t) if x^* is a fixed point of T_t for all $t \in \mathbb{T}$. We say that (T_t) is *asymptotically contractive* on X if $d(T_t x, T_t y) \to 0$ as $t \to \infty$ for all $x, y \in \mathsf{X}$; *order-preserving* if T_t is order preserving on X for all $t \in \mathbb{T}$; and *globally stable* if (T_t) has a unique stationary point $x^* \in \mathsf{X}$ and $d(T_t x, x^*) \to 0$ as $t \to \infty$ for all $x \in \mathsf{X}$. To simplify terminology, in the case where $\mathbb{T} = \mathbb{Z}_+$ and T_t is the t-th composition of $T := T_1$ for all t, we assign properties listed above directly to T. For example, we say that T is *globally stable* if the semigroup $(T_t)_{t\in\mathbb{Z}_+} = (T^t)_{t\in\mathbb{Z}_+}$ is globally stable.

2.3. Results

We can now state our main results in preordered metric space. Throughout, we take $(\mathsf{X}, d, \preceq)$ to be such a space. We use the following assumptions.

Assumption 2.1. The metric d is complete and satisfies the diagonal property.

Assumption 2.2. The preorder $\preceq$ is closed with respect to d.

In this setting we have the following theorem, which is proved in Appendix A.1.

Theorem 2.1. *Let (T_t) be an order-preserving semigroup on X. If Assumptions 2.1 and 2.2 hold, then the following statements are equivalent:*

(i) *(T_t) is asymptotically contractive and has at least one order-bounded trajectory.*

(ii) *(T_t) is globally stable.*

Remark 2.1. Global stability does not imply that every trajectory is order-bounded. For example, let $\mathsf{X} = \mathbb{R}$ with the Euclidean metric and the trivial order ($x \preceq y$ if and only if $x = y$), and let $T_t x = xe^{-t}$. Then Assumptions 2.1 and 2.2 hold and (T_t) is globally stable, but the trajectory of x is order-bounded only when $x = 0$.

Switching to discrete time, we have the following corollary. The corollary is obvious given Theorem 2.1, but nonetheless worth stating due to its independent value as a fixed point theorem.

Corollary 2.1. *Let T be an order-preserving self-map on X. If Assumptions 2.1 and 2.2 hold, then the following statements are equivalent:*

(i) *T is asymptotically contractive and has at least one order-bounded trajectory.*

(ii) *T is globally stable.*

Theorem 2.1 assumes that the preorder $\preceq$ is closed. The next result shows that this assumption can be dropped when (T_t) satisfies additional conditions. The proof is in Appendix A.2.

Theorem 2.2. *If Assumption 2.1 holds and $x \mapsto T_t x$ is continuous on X for all $t \in \mathbb{T}$, then the equivalence of* (i) *and* (ii) *in Theorem 2.1 is also valid.*

3. Stability on Probability Spaces

Now we shift to a setting where X is a family of probability measures on a given set S. (In other words, we consider distribution dynamics under various laws of motion.) We retain this setting for the remainder of the paper. Doing so will allow us to specialize and sharpen our results. One non-trivial change will be reinterpreting order-boundedness of trajectories in terms of tightness.

After developing the necessary background—the Bhattacharya metric on $\mathscr{P}(S)$ and the equivalence between tightness and order boundedness (Proposition 3.1)—we specialize the stability results of Section 2 to semigroups acting on distributions.

3.1. Background

Let $S = (S, \leq)$ be a partially ordered Polish space, that is, S is a separable and completely metrizable topological space equipped with a closed partial order $\leq$. The Borel sets of S are denoted by $\mathscr{B}$. A function $f : S \to \mathbb{R}$ is called *increasing* if $f(x) \leq f(y)$ whenever $x \leq y$. It is called *decreasing* if $f(y) \leq f(x)$ whenever $x \leq y$. This terminology extends to sets in $\mathscr{B}$ by considering their indicator functions. We let bS represent the set of bounded measurable functions from S to $\mathbb{R}$, cbS represent the set of continuous functions in bS, ibS represent the set of increasing functions in bS, and $\mathscr{P}(S)$ be the set of all probability measures on $(S, \mathscr{B})$. Elements of $\mathscr{P}(S)$ are also called *distributions*. For $h \in bS$ and $\phi \in \mathscr{P}(S)$, we set $\phi(h) := \int h d\phi$. We recall that a set $\Phi \subset \mathscr{P}(S)$ is called *tight* if, for all $\epsilon > 0$, there exists a compact $K \subset S$ with $\phi(K^c) \leq \epsilon$ for all $\phi \in \Phi$. For any pair $\phi, \psi \in \mathscr{P}(S)$, write

$$\phi \preceq_{\text{sd}} \psi \quad \iff \quad \phi(h) \leq \psi(h) \text{ for all } h \in ibS.$$

This is the usual notion of (first order) *stochastic dominance*. Under our assumptions on $(S, \leq)$, the relation $\preceq_{\text{sd}}$ is a partial order on $\mathscr{P}(S)$ (see, e.g., [38]). Throughout the paper, we impose the following restriction on S.

Assumption 3.1. A subset of S is compact if and only if it is closed and order bounded.

Here is a simple alternative characterization that is helpful for the proofs.

Lemma 3.1. *Assumption 3.1 holds if and only if order intervals in S are compact and compact subsets of S are order-bounded.*

Proof. Suppose first that Assumption 3.1 holds. Order intervals are closed (since $\leq$ is closed) and order-bounded, hence compact by Assumption 3.1. If K is compact, then K is order-bounded by Assumption 3.1. Conversely, suppose that order intervals are compact and compact sets are order-bounded. If K is compact, then K is closed (since S is Hausdorff) and order-bounded (by assumption). If K is closed and order-bounded, then $K \subseteq [a, b]$ for some $a, b \in S$. Since $[a, b]$ is compact and K is a closed subset, K is compact. □

Example 3.1. Assumption 3.1 holds when S is a product of intervals in $\mathbb{R}^n$ with the componentwise partial order and Euclidean metric. This includes $\mathbb{R}^n$, $[0, \infty)^n$, $(0, \infty)^n$, $(0, 1)^n$, $[0, 1]^n$, and mixed cases such as $[0, 1) \times (0, \infty)$.

The next result will be used to connect tightness with global stability. Its proof is given in Appendix A.3.

Proposition 3.1. *$\Lambda \subset \mathscr{P}(S)$ is tight if and only if Λ is order bounded in $(\mathscr{P}(S), \preceq_{sd})$.*

In this paper, when considering stability of maps over distribution space, we always work with the *Bhattacharya metric* β on $\mathscr{P}(S)$, which is defined by

$$\beta(\phi, \psi) = \sup_{h \in ibS,\ |h| \leq 1} |\phi(h) - \psi(h)|. \tag{3.1}$$

Under Assumption 3.1, the metric β is complete on $\mathscr{P}(S)$ [16, Theorem 4.1].

Remark 3.1. Since $ibS \subset bS$, we have $\beta(\phi, \psi) \leq \|\phi - \psi\|_{TV} := \sup_{h \in bS,\, |h| \leq 1} |\phi(h) - \psi(h)|$, so convergence in total variation implies convergence in β. When $S = \mathbb{R}$ with the usual order, convergence in β is equivalent to uniform convergence of distribution functions, which is strictly weaker than convergence in total variation and strictly stronger than weak convergence. See [16] for further discussion.

Lemma 3.2. *The metric β satisfies the diagonal property on $\mathscr{P}(S)$.*

Proof. Fix ϕ, ψ, ℓ and u in $\mathscr{P}(S)$ with $\phi, \psi \in [\ell, u]$. Given $h \in ibS$ with $|h| \leq 1$, we have $\phi(h) - \psi(h) \leq u(h) - \ell(h)$ and $\psi(h) - \phi(h) \leq u(h) - \ell(h)$. It follows that $|\phi(h) - \psi(h)| \leq u(h) - \ell(h) = |u(h) - \ell(h)| \leq \beta(\ell, u)$. The diagonal property is obtained by taking the sup over $h \in ibS$ with $|h| \leq 1$. □

Another useful observation is as follows.

Lemma 3.3. *The partial order $\preceq_{sd}$ on $\mathscr{P}(S)$ is closed with respect to β.*

Proof. Suppose $(\phi_n), (\psi_n) \subset \mathscr{P}(S)$ with $\phi_n \preceq_{\text{sd}} \psi_n$ for all n, $\beta(\phi_n, \phi) \to 0$, and $\beta(\psi_n, \psi) \to 0$. Fix $h \in ibS$ with $|h| \leq 1$. From convergence in the Bhattacharya metric, we obtain $\phi_n(h) \to \phi(h)$ and $\psi_n(h) \to \psi(h)$ as $n \to \infty$. We also have $\phi_n(h) \leq \psi_n(h)$ for all n. Hence, taking limits, $\phi(h) \leq \psi(h)$. This holds for all $h \in ibS$ with $|h| \leq 1$ and, by scaling, extends to all $h \in ibS$. Therefore $\phi \preceq_{\text{sd}} \psi$. □

3.2. Semigroups over Distributions

Let us now consider a semigroup $(T_t)_{t \in \mathbb{T}}$ on $\mathscr{P}(S)$. The semigroup represents a model that shifts distributions forward in time via $t \mapsto T_t \phi$, where ϕ is some initial distribution. In all of what follows, $\mathscr{P}(S)$ is endowed with the stochastic dominance partial order $\preceq_{\text{sd}}$ and the Bhattacharya metric β. Moreover, $S = (S, \preceq)$ is itself a partially ordered Polish space satisfying Assumption 3.1. In this setting, we can state the following result, which is proved in Appendix A.3.

Theorem 3.1. *Let (T_t) be an order-preserving semigroup on $\mathscr{P}(S)$. The following statements are equivalent:*

(i) *(T_t) is asymptotically contractive and has an order-bounded trajectory in $\mathscr{P}(S)$.*

(ii) *(T_t) is asymptotically contractive and has a tight trajectory in $\mathscr{P}(S)$.*

(iii) *(T_t) is globally stable on $\mathscr{P}(S)$.*

The statement that (T_t) has a tight trajectory means that there exists at least one $\phi \in \mathscr{P}(S)$ such that $(T_t \phi)_{t \in \mathbb{T}}$ is tight in $\mathscr{P}(S)$.

Often the operators that we consider will be Markov operators generated from transition probability functions. Such operators have other useful properties, such as linearity with respect to distributions. We discuss this case in Section 4. We also note, however, that Theorem 3.1 can be used to study nonlinear maps over distributions, which occur frequently in economic dynamics (see, e.g., [19]).

4. Markov Semigroups

In this section we connect the abstract theory to concrete Markov dynamics. The order-preserving semigroups to which Theorem 3.1 applies are, in our applications, generated by stochastic kernels and transition probability functions; here we develop that correspondence and record the properties of the resulting Markov operators.

4.1. Stochastic Kernels

A *stochastic kernel on* S is a map P from $S \times \mathscr{B}$ to $[0,1]$ such that $B \mapsto P(x, B)$ is in $\mathscr{P}(S)$ for each $x \in S$ and $x \mapsto P(x, B)$ is measurable for each $B \in \mathscr{B}$. Each stochastic kernel P on S generates two operators. The first is the *left Markov operator*, defined via

$$\phi \mapsto \phi P, \qquad (\phi P)(B) := \int P(x, B)\phi(\mathrm{d}x) \qquad (B \in \mathscr{B}). \tag{4.1}$$

The second is the *right Markov operator*, defined by

$$h \mapsto Ph, \qquad (Ph)(x) := \int h(y)P(x, \mathrm{d}y) \qquad (h \in bS,\ x \in S). \tag{4.2}$$

In line with much of the literature (see, e.g., [23]), the same symbol P is used for both operators, as well as for the stochastic kernel. In addition, we will drop the terms "left" and "right" when the meaning is clear from context. The map $\phi \mapsto \phi P$ is understood as updating distribution ϕ to the next period, while $(Ph)(x)$ is interpreted as the conditional expectation of $h(X)$ when X is the next period state and x is the current state. For the two operators in (4.1) and (4.2), Fubini's theorem yields the well-known "duality" relation

$$(\phi P)(h) = \phi(Ph) \text{ for any } \phi \in \mathscr{P}(S) \text{ and any } h \in bS. \tag{4.3}$$

The kernel P is called *increasing* if $x \leq x'$ implies $P(x, \cdot) \preceq_{\mathrm{sd}} P(x', \cdot)$. This is equivalent to $Ph \in ibS$ whenever $h \in ibS$, and to $\phi P \preceq_{\mathrm{sd}} \psi P$ whenever $\phi \preceq_{\mathrm{sd}} \psi$. The following result parallels the well-known fact that total variation distance between distributions is not increased after passing them through a Markov kernel.

Lemma 4.1. *If P is an increasing kernel, then $\beta(\phi P, \psi P) \leq \beta(\phi, \psi)$ for all $\phi, \psi \in \mathscr{P}(S)$.*

Proof. Observe that, for probability measures $\phi, \psi \in \mathscr{P}(S)$, we have

$$\beta(\phi P, \psi P) = \sup_h |(\phi P)(h) - (\psi P)(h)| = \sup_h |\phi(Ph) - \psi(Ph)|,$$

where the supremum is over $h \in ibS$ with $|h| \leq 1$ and the second equality is by the duality relationship in (4.3). Since P is increasing and h is increasing, Ph is increasing. Since $|h| \leq 1$, we have $|Ph| \leq 1$, so the last expression is dominated by $\beta(\phi, \psi)$. □

4.2. Transition Probability Functions

Following standard terminology, a *transition probability function* is a family of stochastic kernels $(P_t)_{t \in \mathbb{T}}$ such that $P_0(x, \cdot) = \delta_x$ for all $x \in S$ and the Chapman–Kolmogorov equation holds:

$$P_{s+t}(x, B) = \int P_t(y, B)P_s(x, \mathrm{d}y) \qquad (s, t \in \mathbb{T},\ x \in S,\ B \in \mathscr{B}). \tag{4.4}$$

When $(P_t)_{t \in \mathbb{T}}$ is a transition probability function, the family of maps $\phi \mapsto \phi P_t$,

$$(\phi P_t)(B) := \int P_t(x, B)\phi(\mathrm{d}x) \qquad (B \in \mathscr{B},\ t \in \mathbb{T}) \tag{4.5}$$

forms a semigroup on $\mathscr{P}(S)$, with the semigroup property $P_{s+t} = P_s \circ P_t$ following from the Chapman–Kolmogorov relations. In this setting, we say that $(P_t)_{t \in \mathbb{T}}$ is a *Markov semigroup on* S. We use the same symbol $(P_t)_{t \in \mathbb{T}}$ to represent the Markov semigroup and the transition probability function, since both represent the same object.[1]

Remark 4.1. Each transition probability function $(P_t)_{t \in \mathbb{T}}$ also defines a family of operators $h \mapsto P_t h$, indexed by $t \in \mathbb{T}$, each P_t mapping bS to bS, through the right Markov operator relation in (4.2). Some authors have this in mind when they refer to the semigroup generated by the transition probability function, rather than the family of maps $\phi \mapsto \phi P_t$ sending $\mathscr{P}(S)$ to $\mathscr{P}(S)$. This is just a matter of convention. Our terminology is similar to [39] and [40].

[1] The transition probability function defines the Markov semigroup on S via (4.5). The semigroup defines the transition probability function via $P_t(x, B) := (\delta_x P_t)(B)$ for all $x \in S$, $t \in \mathbb{T}$ and $B \in \mathscr{B}$.

Example 4.1. In the discrete time case, where we work with a fixed stochastic kernel P on S, we set $P_0 := I$ and $P_t := P^t$ for all $t \in \mathbb{N}$. The family $(P_t)_{t \in \mathbb{Z}_+}$ is a transition probability function.

Example 4.2. Let $S = \mathbb{R}$ and $P(t, x, \mathrm{d}y) = p(t, x, y)\,\mathrm{d}y$, where

$$p(t, x, y) := \sqrt{\frac{\theta}{\pi\sigma^2(1 - e^{-2\theta t})}} \exp\left(-\frac{\theta(y - xe^{-\theta t})^2}{\sigma^2(1 - e^{-2\theta t})}\right).$$

This is the transition density function for the Ornstein–Uhlenbeck process

$$dX_t = -\theta X_t\,dt + \sigma\,dW_t, \quad X_0 = x, \tag{4.6}$$

where $\theta > 0$ is the mean-reversion rate, $\sigma > 0$ is the volatility parameter, and $(W_t)_{t \geq 0}$ is Brownian motion. The family of maps $(P_t)_{t \geq 0}$ defined by $\phi \mapsto \phi P_t$ with $(\phi P_t)(B) := \int \int_B p(t, x, y)\ \mathrm{d}y\,\phi(\mathrm{d}x)$ is the Markov semigroup on $\mathbb{R}$ generated by the Ornstein–Uhlenbeck transition probability function. For (4.6), the solution can be expressed as

$$X_t = e^{-\theta t}x + \sigma_t Z_t, \quad \text{where} \quad \sigma_t := \sigma\sqrt{\frac{1 - e^{-2\theta t}}{2\theta}} \tag{4.7}$$

and Z_t is a standard normal random variable for each t. Hence, for the right Markov operators, we have

$$(P_t h)(x) = \mathbb{E}[h(e^{-\theta t}x + \sigma_t Z)] \qquad (x \in S,\ h \in bS), \tag{4.8}$$

where Z is standard normal.

A transition probability function $(P_t)_{t \in \mathbb{T}} = (P_t)$ will be called *increasing* if P_t is an increasing stochastic kernel for all $t \in \mathbb{T}$. For example, (4.8) implies that $P_t h$ is increasing whenever h is increasing, so the Ornstein–Uhlenbeck transition probability function is increasing. Evidently, $(P_t)_{t \in \mathbb{T}}$ is an increasing transition probability function if and only if it is order-preserving when viewed as a semigroup on $(\mathscr{P}(S), \preceq_{\mathrm{sd}})$, as in (4.5).

We say that a transition probability function (P_t) is *bounded in probability* when the family of distributions $(P_t(x, \cdot))_{t \in \mathbb{T}}$ is tight for all $x \in S$. For example, for the Ornstein–Uhlenbeck process, the representation (4.7) and $\sigma_t \leq \bar{\sigma} := \sigma/\sqrt{2\theta}$ for all t give $|X_t| \leq |x| + \bar{\sigma}|Z_t|$. Since each Z_t is standard normal, the family $(P_t(x, \cdot))_{t \geq 0}$ is tight.

We recall that an S-valued stochastic process $(X_t) := (X_t)_{t \in \mathbb{T}}$ supported on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *Markov process* if

$$\mathbb{P}\{X_t \in B \mid \mathcal{F}_s\} = \mathbb{P}\{X_t \in B \mid X_s\} \ \ \mathbb{P}\text{-a.s. for all } s, t \in \mathbb{T} \text{ with } s \leq t \text{ and all } B \in \mathscr{B},$$

where $(\mathcal{F}_t)_{t \in \mathbb{T}}$ is a filtration to which (X_t) is adapted. Unless otherwise stated, $(\mathcal{F}_t)$ is the natural filtration generated by (X_t); otherwise we say that (X_t) is Markov *with respect to* $(\mathcal{F}_t)$.

Given a transition probability function $(P_t)_{t \in \mathbb{T}}$, we say that a Markov process $(X_t)_{t \in \mathbb{T}}$ is $(P_t)_{t \in \mathbb{T}}$*-Markov* when

$$\mathbb{P}\{X_t \in B \mid \mathcal{F}_s\} = P_{t-s}(X_s, B) \quad \text{for all nonnegative } s \leq t \text{ and all } B \in \mathscr{B}.$$

For the case of discrete time, when P_t is just the t-composition P^t for some fixed stochastic kernel P, we will simply say that (X_t) is P*-Markov*. In other words, $(X_t)_{t \in \mathbb{Z}_+}$ is P-Markov when

$$\mathbb{P}\{X_{t+1} \in B \mid \mathcal{F}_t\} = P(X_t, B) \quad \text{for all } t \in \mathbb{Z}_+ \text{ and all } B \in \mathscr{B}.$$

Given $x \in S$, we write $\mathbb{P}_x$ for the conditional probability $\mathbb{P}\{\cdot \mid X_0 = x\}$ and $\mathbb{E}_x$ for the corresponding expectation. For a pair of processes $((X_t), (X'_t))$, $\mathbb{P}_{x,x'}$ and $\mathbb{E}_{x,x'}$ are defined analogously, conditioning on $X_0 = x$ and $X'_0 = x'$.

5. Markov Stability

The characterization in Theorem 3.1 reduces global stability to asymptotic contractivity together with a tight trajectory. We now develop conditions under which these hold for monotone Markov processes. Since the state space is not required to be compact or order bounded, we need some restriction on divergence, so that probability mass does not vanish towards the "edges" of the state space. Here this takes the form of a tightness condition, connected to order boundedness of trajectories via Proposition 3.1. We treat discrete time in Section 5.1 and continuous time in Section 5.2. In Section 5.3 we strengthen the assumptions on the state space to obtain a uniform exponential rate of convergence.

5.1. Discrete Time Theory

As before, we take $(S, \leq)$ to be a partially ordered Polish space and $\preceq_{\mathrm{sd}}$ to be stochastic dominance on $\mathscr{P}(S)$. The space $(S, \leq)$ is assumed to satisfy Assumption 3.1. A distribution ϕ is called *excessive* for a stochastic kernel P on S if $\phi P \preceq_{\mathrm{sd}} \phi$ and *deficient* for P if $\phi \preceq_{\mathrm{sd}} \phi P$. As usual, P is called *Feller* if $Ph \in cbS$ whenever $h \in cbS$.

Let P be a stochastic kernel on S. As in [13], we say that P is *order reversing* if there exist two independent P-Markov processes $(X_t)_{t \in \mathbb{Z}_+}$ and $(X'_t)_{t \in \mathbb{Z}_+}$ on a common probability space such that, for every $x, x' \in S$ with $x' \leq x$, there exists a $t \in \mathbb{Z}_+$ with $\mathbb{P}_{x,x'}\{X_t \leq X'_t\} > 0$. Intuitively, order reversing means that no initial ordering is permanent: regardless of how the two chains start, randomness can eventually (weakly) reverse their positions. We can now state our main discrete time result.

Theorem 5.1. *Let P be an increasing stochastic kernel on S. If P is order reversing, then P is globally stable if and only if P is bounded in probability.*

The key idea behind the proof of Theorem 5.1 (given in Appendix A.5) is that boundedness in probability, combined with order reversing, provides a form of monotone mixing that leads to asymptotic contractivity. Crucially, the asymptotic contractivity is with respect to a metric defined via monotone test functions (the Bhattacharya metric), so a monotone form of mixing suffices—we do not need anything as strong as irreducibility. Global stability then follows from Theorem 3.1, since boundedness in probability implies the existence of a tight trajectory.

To place Theorem 5.1 in context, it is helpful to compare it to Theorems 1 and 2 of [13], where it is shown that an increasing and order reversing stochastic kernel P on S is globally stable whenever

(KS1) P is bounded in probability and has an excessive or a deficient distribution, or

(KS2) P is bounded in probability and Feller.

Theorem 5.1 improves both (KS1) and (KS2). In particular, it shows that boundedness in probability is enough for global stability when P is increasing and order reversing. We do not need to separately check for the existence of an excessive or deficient distribution, as in (KS1), nor do we require the Feller condition, as in (KS2).

Theorem 5.1 implies an ergodicity result: If P is increasing, order reversing, and is bounded in probability, then P has a unique stationary distribution ϕ^* and, for every P-Markov process (X_t), $\lim_{n\to\infty} \frac{1}{n}\sum_{t=0}^{n-1} h(X_t) = \phi^*(h)$ with probability one for all $h \in ibS$, and also for all $h \in cbS$. This follows from Theorem 5.1 and Proposition 4.1 in [14].

5.2. Continuous Time Theory

We now turn to continuous time. The mixing conditions in this section serve to deliver the asymptotic contractivity required by Theorem 3.1. The mechanism is as follows. If two copies of a monotone process, started at x and x', become ordered at some point, then monotonicity implies that, from that point on, the conditional distribution of the first given the history is dominated by that of the second. If ordering occurs with probability one, then applying the same argument with the roles of x and x' reversed shows that the two distributions are drawn together in the Bhattacharya metric.

Theorem 5.1 relies on independence: order reversing only requires a positive probability that two independent chains reverse their order, and independence, combined with boundedness in probability, upgrades this to ordering with probability one [13, Lemma A.5]. In continuous time, independence is less convenient. For example, two independent copies of the PDMPs in Section 6 jump at different times, so their deterministic flows run for different lengths of time between jumps, which makes the ordering hard to control. It is more natural to let the copies jump at the same times. We therefore allow the two processes to be dependent and, in exchange, require that they become ordered with probability one.

Formally, we call a stochastic kernel P on S *weakly order mixing* if there exists an $(S \times S)$-valued process $((X_k), (X'_k))_{k \in \mathbb{Z}_+}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, adapted to a filtration $(\mathcal{F}_k)$, such that (X_k) and (X'_k) are both P-Markov with respect to $(\mathcal{F}_k)$ and

$$\mathbb{P}_{x,x'}\{X_k \leq X'_k \text{ for some } k \in \mathbb{Z}_+\} = 1 \quad \text{for all } x, x' \in S.$$

We extend this notion to a transition probability function (P_t) on S by calling (P_t) *weakly order mixing* if P_u is weakly order mixing for some $u > 0$.

In continuous time, weak order mixing can typically be verified as follows. Since the definition only observes the processes at times ku, it is convenient to ensure that, once ordered, the two copies stay ordered. Two copies of the process are run until they first become ordered, sharing some sources of randomness (such as jump times) but not others (such as the sizes of the jumps). Both copies are then "restarted" from their current states, now sharing all sources of randomness, which preserves the ordering when the dynamics are monotone for each realization of the randomness, as in Assumption 6.1. The proof of Lemma A.8 follows this pattern.

We can now state our main continuous time result for a transition probability function (P_t) on S.

Theorem 5.2. *Let (P_t) be increasing. If (P_t) is weakly order mixing, then (P_t) is globally stable if and only if (P_t) has at least one tight trajectory.*

Compared with Theorem 5.1, boundedness in probability is replaced by the existence of a single tight trajectory. This is possible because weak order mixing already requires ordering with probability one, so no additional argument is needed to obtain it. The proof is given in Appendix A.5. We apply this result to piecewise deterministic Markov processes in Section 6 below.

5.3. Uniform Exponential Convergence

The characterizations obtained so far establish global stability but provide no rate of convergence to the stationary distribution. By strengthening the assumptions on the state space, we can say more. In this subsection we suppose that the state space is compact with least and greatest elements—the classical monotone mixing setting—and show that, under a monotone mixing condition, the process converges to its stationary distribution at a uniform exponential rate.

This recovers and sharpens the classical monotone mixing result for monotone Markov chains [11]: we additionally (a) establish an explicit exponential rate of convergence, (b) prove ergodicity for bounded increasing observables, and (c) extend the framework from discrete to continuous time.

As before $(S, \leq)$ is a partially ordered Polish space and $\preceq_{\text{sd}}$ is stochastic dominance. Following [11], we suppose that S is compact and, in addition, has a greatest element b and a least element a. Note that, under these conditions, Assumption 3.1 is satisfied. Given $x \in S$, we write U_x for all $y \in S$ with $x \leq y$ and D_x for all $y \in S$ with $y \leq x$.

In line with [11], we say that a stochastic kernel P on S satisfies the *monotone mixing condition* (MMC) if there exist an $\hat{x} \in S$ and an $\epsilon > 0$ such that

$$P(a, U_{\hat{x}}) \geq \epsilon \quad \text{and} \quad P(b, D_{\hat{x}}) \geq \epsilon. \tag{5.1}$$

In other words, a chain starting at a rises above $\hat{x}$ with positive probability, while a chain starting from b falls below $\hat{x}$ with positive probability.

Remark 5.1. Let $S = [0, 1]$ and $X_{t+1} = \rho X_t + (1 - \rho) W_{t+1}$, where $\rho \in [0, 1)$ and (W_t) is IID and uniform on $[0, 1]$. The kernel P is increasing. The support of $P^n(0, \cdot)$ is $[0, 1 - \rho^n]$ and that of $P^n(1, \cdot)$ is $[\rho^n, 1]$, so P^n satisfies the MMC with $\hat{x} = 1/2$ whenever $\rho^n < 1/2$. For $\rho \geq 1/2$, the MMC fails for $n = 1$, which is why Theorem 5.3 requires it only for some P_u. For economic applications, see [11] and Chapter 12 of [41].

We now present the main result of this subsection, which extends Theorem 2 of [11]. In the statement, S has the properties listed at the start of this subsection.

Theorem 5.3. *Let $(P_t)_{t \in \mathbb{T}}$ be a transition probability function on S. If $(P_t)_{t \in \mathbb{T}}$ is increasing and there exists a positive $u \in \mathbb{T}$ such that P_u satisfies the MMC, then $(P_t)_{t \in \mathbb{T}}$ is globally stable on $\mathscr{P}(S)$ under β and convergence to the unique stationary distribution ϕ^* is exponential:*

$$\beta(\phi P_t, \phi^*) \leq C e^{-\alpha t} \quad \text{for all } \phi \in \mathscr{P}(S) \text{ and } t \in \mathbb{T}, \tag{5.2}$$

where ϵ is the constant in (5.1) for P_u, $C := 2/(1-\epsilon)$ and $\alpha := \ln(1/(1-\epsilon))/u$. Moreover, when $\mathbb{T} = \mathbb{Z}_+$, each P-Markov process (X_t) is monotone ergodic, with $\lim_{n \to \infty} \frac{1}{n} \sum_{t=0}^{n-1} h(X_t) = \phi^*(h)$ with probability one for all $h \in ibS$.

The proof can be found in Appendix A.7.

6. Piecewise Deterministic Markov Processes

Building on the continuous time theory of Section 5.2, we develop in this section a general stability result for piecewise deterministic Markov processes (PDMPs) that can be applied off the shelf in a range of settings. It is applied to a model of income dynamics with deterministic drift in Section 7.2. For existence and uniqueness of stationary distributions of PDMPs via the embedded post-jump chain, see, e.g., [42, 43].

Let $(S, \leq)$ be a partially ordered Polish space satisfying Assumption 3.1. A *piecewise deterministic Markov process* (PDMP) on S is generated by the following characteristics:

(i) a *semi-flow* $\Phi : S \times \mathbb{R}_+ \to S$,

(ii) a *jump intensity* $\lambda \in (0, \infty)$,

(iii) a *jump shock space* U and shock distribution μ on U, and

(iv) a *jump function* $F : S \times U \to S$.

Here U is a Polish space and the jump function F is assumed to be measurable. The map Φ is called a *semi-flow* because it satisfies $\Phi(x, 0) = x$ and $\Phi(\Phi(x, s), t) = \Phi(x, s + t)$ for all $x \in S$ and $s, t \geq 0$. We assume that Φ is continuous. Informally, the process evolves as follows:

(i) Follow the deterministic flow: $X_t = \Phi(X_0, t)$ until the first jump.

(ii) At the first jump time T, draw ξ from μ and set $X_T = F(X_{T-}, \xi)$.

(iii) Repeat from the new state X_T.

In most applications, the semi-flow is the solution to an ODE. For example, if $S = \mathbb{R}$ and the deterministic dynamics have the linear form $\dot{x} = ax$, then $\Phi(x, t) = x \exp(at)$.

To formalize the construction of the associated Markov process $(X_t)_{t \geq 0}$, we first collect the following components: an S-valued random element X_0 with distribution $\psi \in \mathscr{P}(S)$, a real-valued IID sequence $(E_n)_{n \in \mathbb{N}}$ with common distribution $\text{Exp}(\lambda)$, and a U-valued IID sequence $(\xi_n)_{n \in \mathbb{N}}$ with common distribution μ. All random elements live on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We first build the *embedded chain* $(Z_n)_{n \in \mathbb{Z}_+}$ via

$$Z_0 = X_0 \quad \text{and} \quad Z_{n+1} = F(\Phi(Z_n, E_{n+1}), \xi_{n+1}).$$

We define the jump times $T_0 := 0$ and $T_n := \sum_{i=1}^n E_i$, while $N_t := \max\{n \geq 0 \,:\, T_n \leq t\}$ counts the number of jumps up until time t. Finally, we construct $(X_t)_{t \geq 0}$ via $X_t = \Phi(Z_{N_t}, t - T_{N_t})$ for all $t \geq 0$. A more intuitive representation is

$$X_t = \Phi(Z_n, t - T_n) \quad \text{for } t \in [T_n, T_{n+1}) \text{ and } n \in \mathbb{Z}_+.$$

Sample paths are càdlàg, with smooth deterministic segments broken by jumps. The process (X_t) is a Markov process [40, Theorem 2.4]. We work with the following assumptions.

Assumption 6.1 (PDMP Monotonicity)**.** The primitives are such that

(i) For all $t \geq 0$, the function $x \mapsto \Phi(x, t)$ is order preserving on S.

(ii) For all $u \in U$, the function $x \mapsto F(x, u)$ is order preserving on S.

In the next assumption, ξ and ξ' are independent draws from μ.

Assumption 6.2 (PDMP Mixing)**.** There exist measurable functions $f_1, f_2 : U \to S$ such that

$$f_1(\xi) \leq F(x, \xi) \leq f_2(\xi) \text{ for all } x \in S \text{ and } \mathbb{P}\{f_2(\xi) \leq f_1(\xi')\} > 0.$$

The transition probability function corresponding to the PDMP $(X_t)_{t \geq 0}$ is given by

$$P_t(x, B) = \mathbb{P}\{X_t \in B \mid X_0 = x\} \qquad (t \in \mathbb{R}_+,\ x \in S,\ B \in \mathscr{B}).$$

Under the stated monotonicity and mixing conditions, we can prove the following result.

Theorem 6.1. *If Assumptions 6.1 and 6.2 hold, then the PDMP is globally stable.*

The proof is based on Theorem 5.2 and can be found in Appendix A.6. The strategy is to verify the three hypotheses of Theorem 5.2. Monotonicity is established by coupling two copies of the PDMP with the same jump times and the same shocks: if one copy starts above the other, then Assumption 6.1 ensures that it stays above through both the deterministic flow and the jumps. Boundedness in probability follows from the pathwise bounds in Assumption 6.2, which prevent post-jump states from escaping to infinity, while the exponential inter-arrival times prevent the deterministic flow from carrying the process too far between jumps. The most interesting step is weak order mixing, which uses a different coupling: the two copies share jump times but receive independent shocks. The condition $\mathbb{P}\{f_2(\xi) \leq f_1(\xi')\} > 0$ in Assumption 6.2 then guarantees that, at each shared jump, there is a positive probability that the two processes become ordered, regardless of their current states. From then on, the copies share shocks as well, so the ordering is preserved.

Now we turn to applications.

7. Applications

We now illustrate the theory with two applications.

7.1. Learning with Belief Shocks

Belief shocks are used in macroeconomics and finance to generate effects such as countercyclical uncertainty and forecast biases (see, e.g., [31, 32, 44]). Here we use Theorem 5.1 to study a model of Bayesian learning subject to belief shocks. In the model, an agent faces uncertainty about a fixed hidden state $\theta \in \{\ell, h\}$. Let $\pi_t \in (0, 1)$ represent the agent's belief that $\theta = h$ at time t. The state π_t is driven by Bayesian updating while observing IID signals $(Z_t)_{t\in\mathbb{N}}$, drawn from density f_h if $\theta = h$ and f_ℓ if $\theta = \ell$. Let $L(z) := f_h(z)/f_\ell(z)$ denote the likelihood ratio. In terms of the log-odds $\eta_t := \ln(\pi_t/(1-\pi_t))$, Bayes' rule becomes

$$\eta_{t+1} = \eta_t + \xi_{t+1}, \quad \text{where} \quad \xi_{t+1} := \ln L(Z_{t+1}). \tag{7.1}$$

The state space for $(\eta_t)_{t\in\mathbb{Z}_+}$ is $\mathbb{R}$.

We introduce occasional belief shocks that reset the agent's posterior according to a stochastic kernel Q on $\mathbb{R}$. Under these shocks, the belief process $(\eta_t)_{t\in\mathbb{Z}_+}$ evolves on $\mathbb{R}$ according to

$$\eta_{t+1} = I_{t+1} \cdot R_{t+1} + (1 - I_{t+1}) \cdot (\eta_t + \xi_{t+1}), \tag{7.2}$$

where $(I_t)_{t\in\mathbb{N}}$ is the belief reset indicator process and $R_{t+1} \sim Q(\eta_t, \cdot)$ is the reset value. The stochastic kernel Q is assumed to be increasing. Intuitively, this means that beliefs have some positive correlation over the reset. It includes the special case where reset values are IID. The reset indicator process is IID and each I_t is drawn from Bernoulli(ρ).

In what follows, we specialize to $\theta = h$, so that each Z_t is drawn from f_h. (Analysis under $\theta = \ell$ is very similar.) Let P denote the stochastic kernel corresponding to (7.2) under this specialization.

Lemma 7.1. *The stochastic kernel P for the belief process is increasing.*

Proof. Take $\eta_1 \leq \eta_2$. Since Q is increasing, $Q(\eta_1, \cdot) \preceq_{\mathrm{sd}} Q(\eta_2, \cdot)$. By Strassen's theorem (see, e.g., [3]), there exists a pair of random variables R_1 and R_2 such that the first component R_1 is drawn from $Q(\eta_1, \cdot)$, the second component R_2 is drawn from $Q(\eta_2, \cdot)$, and $R_1 \leq R_2$ almost surely. Let I and ξ be drawn from their respective distributions, so that

$$\eta_i' := I \cdot R_i + (1 - I) \cdot (\eta_i + \xi) \text{ has distribution } P(\eta_i, \cdot) \text{ for } i = 1, 2.$$

If $I = 0$, then $\eta_1' = \eta_1 + \xi \leq \eta_2 + \xi = \eta_2'$. If $I = 1$, then $\eta_1' = R_1$ and $\eta_2' = R_2$. In both cases, $\eta_1' \leq \eta_2'$. As a result, for any $h \in ibS$, $(Ph)(\eta_1) = \mathbb{E}h(\eta_1') \leq \mathbb{E}h(\eta_2') = (Ph)(\eta_2)$. This proves that P is increasing. □

We impose the following conditions:

(A1) The log-likelihood ratio has finite variance: $\sigma^2 = \mathrm{Var}(\ln L(Z)) < \infty$.

(A2) For all $\varepsilon > 0$, there exists a compact $K \subset \mathbb{R}$ such that $\inf_{x\in\mathbb{R}} Q(x, K) \geq 1 - \varepsilon$.

(A3) For every $M > 0$, we have $\mathbb{P}\{L(Z) > M\} > 0$ when $Z \sim f_h$.

Assumptions (A1) and (A3) hold, for example, when f_h and f_ℓ are Gaussian with a common variance. Assumption (A2) is a tightness condition on the reset kernel and holds trivially when resets are IID.

Lemma 7.2. *Under (A1) and (A2), P is bounded in probability.*

Proof. Fix $x \in \mathbb{R}$ and $\varepsilon > 0$. By (A2), choose $K = [a, b]$ such that $Q(y, K) \geq 1 - \varepsilon/3$ for all y. Choose T large enough that $(1-\rho)^T < \varepsilon/3$. For $t \geq T$, let N_t denote the number of updates since the last reset prior to t. The probability that no reset has occurred by time t is $(1-\rho)^t \leq (1-\rho)^T < \varepsilon/3$. Conditional on a reset occurring, the post-reset state lies in K with probability at least $1-\varepsilon/3$. Conditional on $N_t = n$ and the last reset landing in K, we have $\eta_t \in [a+S_n, b+S_n]$ where $S_n = \xi_1 + \cdots + \xi_n$. Let $m := \mathbb{E}\xi_1$, which is finite by (A1). The random variable N_t is stochastically dominated by $N \sim \text{Geometric}(\rho)$. By Chebyshev's inequality, $\mathbb{P}\{|S_n| > c\} \leq \mathbb{1}\{n|m| > c/2\} + 4n\sigma^2/c^2$, which is increasing in n. Choose c large enough that

$$\mathbb{P}\{N|m| > c/2\} + \sum_{n=0}^{\infty} \rho(1-\rho)^n \cdot \frac{4n\sigma^2}{c^2} < \frac{\varepsilon}{3}.$$

Then $\mathbb{P}_x\{\eta_t \notin [a-c, b+c]\} < \varepsilon$ for all $t \geq T$. For $t < T$, each $P^t(x, \cdot)$ is a single distribution and hence tight. A finite union of compact sets covers these, completing the proof. ☐

Lemma 7.3. *Under* (A3)*, the kernel P is order reversing.*

Proof. Let (η_t) and (η_t') be independent P-Markov processes with initial conditions $\eta > \eta'$. We will show that $\mathbb{P}\{\eta_1 < \eta_1'\} > 0$. To this end, let (I, R, ξ) and (I', R', ξ') be the independent shocks for the two processes, where $R \sim Q(\eta, \cdot)$ and $R' \sim Q(\eta', \cdot)$. Consider the event $A = \{I = 1,\ I' = 0,\ \xi' > R - \eta'\}$. On A, process 1 resets to $\eta_1 = R$ while process 2 updates to $\eta_1' = \eta' + \xi' > R = \eta_1$, so $\eta_1 < \eta_1'$. Since the two processes are independent, I, R, and ξ' are mutually independent. We have $\mathbb{P}\{I = 1\} = \rho > 0$, $\mathbb{P}\{I' = 0\} = 1 - \rho > 0$, and, for every realization $R = r$, $\mathbb{P}\{\xi' > r - \eta'\} = \mathbb{P}\{L(Z) > e^{r-\eta'}\} > 0$ by (A3). Hence $\mathbb{P}(A) > 0$, proving that P is order reversing. ☐

Combining the preceding lemmas with Theorem 5.1, we obtain global stability.

Proposition 7.1. *Under* (A1)–(A3)*, the belief shock kernel P is globally stable on $(\mathscr{P}(\mathbb{R}), \beta)$.*

Proof. The state space $S = \mathbb{R}$ satisfies Assumption 3.1 since order intervals $[a, b]$ are compact and compact sets are bounded, hence order-bounded. By Lemma 7.2, P is bounded in probability. By Lemma 7.1, P is increasing. By Lemma 7.3, P is order reversing. Therefore, by Theorem 5.1, P is globally stable. ☐

Figure 2 shows a simulation of beliefs π_t, with Gaussian signals with unit variance and means 0.3 under h and 0 under ℓ, reset probability $\rho = 0.04$, and IID $N(0, 0.25)$ resets in log-odds space. Gradual learning is interrupted by resets toward $\pi = 0.5$; convergence of the empirical distribution, computed from 200,000 periods, to the stationary distribution follows from the ergodicity result implied by Theorem 5.1.

7.2. Income Dynamics with Deterministic Drift

We consider a model of income dynamics with a continuous drift component and random resets, related to the diffusion-based specification in [35]. In contrast to that specification, we avoid the assumption of positive quadratic variation (e.g., Brownian motion), which is difficult to motivate from data. We analyze the model using Theorem 6.1.

Log income X_t is a PDMP on $S = \mathbb{R}$, as in Section 6, with semi-flow Φ generated by the ODE $\dot{x} = g(x)$, where g is locally Lipschitz with at most linear growth, jump intensity $\lambda > 0$, shock space $U = \mathbb{R}$ with shock distribution ν, and jump function $F(x, \zeta) = h(x) + \zeta$, where $h \in ibS$. Thus, log income drifts deterministically between jumps and, at each jump, moves from x to $h(x) + \zeta$.

The next proposition gives a simple sufficient condition for global stability.

Proposition 7.2. *If the support of ν contains points a, b with $b - a > \sup h - \inf h$, then the log income process is globally stable: there exists a unique stationary distribution ϕ^* and*

$$\lim_{t\to\infty} \beta(\phi P_t, \phi^*) = 0 \quad \textit{for all} \quad \phi \in \mathscr{P}(\mathbb{R}). \tag{7.3}$$

Proof. We verify Assumptions 6.1 and 6.2. Then we apply Theorem 6.1.

Regarding Assumption 6.1 (i), since g is locally Lipschitz, solutions of $\dot{x} = g(x)$ are unique forward and backward in time, so distinct trajectories never meet. By continuity of Φ and the intermediate value theorem, $x_1 < x_2$ therefore implies $\Phi(x_1, t) < \Phi(x_2, t)$ for all $t \geq 0$. Regarding Assumption 6.1 (ii), for each $\zeta \in \mathbb{R}$, the map $x \mapsto F(x, \zeta) = h(x) + \zeta$ is increasing since h is assumed to be increasing. Regarding Assumption 6.2, set $f_1(\zeta) = \underline{h} + \zeta$ and $f_2(\zeta) = \bar{h} + \zeta$, where $\underline{h} := \inf h$ and $\bar{h} := \sup h$. By construction, $f_1(\zeta) \leq F(x, \zeta) \leq f_2(\zeta)$ for all $x \in \mathbb{R}$. For the coupling condition, let ζ, ζ' be independent draws from ν. Then $\mathbb{P}\{f_2(\zeta) \leq f_1(\zeta')\} = \mathbb{P}\{\zeta' - \zeta \geq \bar{h} - \underline{h}\} > 0$, where the strict inequality follows from our assumption on the support of ν. Global stability now follows from Theorem 6.1. □

When the drift rate g is constant and resets are pure (h constant), the stationary distribution of income $Y = e^X$ has a Pareto tail with exponent λ/g, determined by the balance between deterministic growth and random resetting. An analogous result for a closely related pure-jump model is proved in [45].

8. Conclusion

We have shown that an order-preserving semigroup on a complete preordered metric space is globally stable if and only if it is asymptotically contractive and has at least one order-bounded trajectory. When applied to monotone Markov models on partially ordered Polish spaces, this yields a necessary and sufficient condition for the existence, uniqueness, and global stability of stationary distributions. The mixing conditions used in this paper could potentially be weakened further, for example by allowing state-dependent jump intensities in the PDMP framework. The Pareto tail result for income discussed in Section 7.2 is specialized and may hold under weaker assumptions.

A. Proofs

This appendix collects remaining proofs.

A.1. Proof of Theorem 2.1

Throughout this section, Assumptions 2.1 and 2.2 are in force.

Regarding (ii) implies (i), let (T_t) be globally stable with stationary point $x^* \in \mathsf{X}$. Then the sequence given by $T_t x^*$ for all $t \in \mathbb{T}$ is an order-bounded trajectory of (T_t). Moreover, picking arbitrary $x, y \in \mathsf{X}$ and applying the triangle inequality,

$$d(T_t x, T_t y) \leq d(T_t x, x^*) + d(x^*, T_t y) \to 0 \qquad (t \to \infty),$$

where the convergence holds by global stability. Hence (T_t) is asymptotically contractive.

For (i) implies (ii), we assume that (T_t) is asymptotically contractive on X with at least one order-bounded trajectory. Due to the last property, we can take $a, b, \bar{x} \in \mathsf{X}$ such that $a \preceq T_t\, \bar{x} \preceq b$ for all $t \in \mathbb{T}$.

Lemma A.1. *For all $\epsilon > 0$, there exists a $\tau \in \mathbb{T}$ such that*

$$s, t \in \mathbb{T} \text{ with } \tau \leq s, t \quad \implies \quad d(T_s\, \bar{x}, T_t\, \bar{x}) < \epsilon. \tag{A.1}$$

Proof. Fix $\epsilon > 0$ and use asymptotic contractivity to choose $\tau \in \mathbb{T}$ such that $d(T_\tau\, a, T_\tau\, b) < \epsilon$. For $s, t \geq \tau$ we have

$$d(T_s\, \bar{x}, T_t\, \bar{x}) = d(T_\tau T_{s-\tau}\, \bar{x}, T_\tau T_{t-\tau}\, \bar{x}) \leq d(T_\tau\, a, T_\tau\, b) < \epsilon.$$

The first inequality is due to the diagonal property of d, combined with the fact that $T_{s-\tau}\, \bar{x}$ and $T_{t-\tau}\, \bar{x}$ both lie in $[a, b]$, which in turn implies that $T_\tau\, a \preceq T_\tau T_{s-\tau}\, \bar{x} \preceq T_\tau\, b$ and $T_\tau\, a \preceq T_\tau T_{t-\tau}\, \bar{x} \preceq T_\tau\, b$. □

Lemma A.2. *There exists an $x^* \in \mathsf{X}$ such that*

$$d(T_t\, \bar{x}, x^*) \to 0 \quad \textit{as} \quad t \to \infty. \tag{A.2}$$

Proof. Let $(t_n)_{n \in \mathbb{N}}$ be an increasing sequence in $\mathbb{T}$ with $t_n \to \infty$ as $n \to \infty$. Consider the sequence defined by $x_n := T_{t_n}\, \bar{x}$. This sequence is Cauchy because, given $\epsilon > 0$, we can choose $\tau \in \mathbb{T}$ such that (A.1) holds and then $d(x_n, x_m) < \epsilon$ for all n, m such that $t_n, t_m \geq \tau$. Since X is complete, it follows that (x_n) converges to a limit x^*. That is,

$$d(T_{t_n}\, \bar{x}, x^*) \to 0 \quad \text{as} \quad n \to \infty. \tag{A.3}$$

We claim that this can be extended to (A.2). To see this, fix $\epsilon > 0$ and choose τ as in (A.1). For $t \geq \tau$ and $n \in \mathbb{N}$ such that $t_n \geq \tau$, we have

$$d(T_t\, \bar{x}, x^*) \leq d(T_t\, \bar{x}, T_{t_n}\, \bar{x}) + d(T_{t_n}\, \bar{x}, x^*) < \epsilon + d(T_{t_n}\, \bar{x}, x^*).$$

Applying (A.3) and taking the limit in n completes the argument. □

Lemma A.3. *The point x^* in Lemma A.2 is stationary for (T_t).*

Proof. Fix $s, t \in \mathbb{T}$. From the existence of an order-bounded trajectory we have $a \preceq T_t\, \bar{x} \preceq b$. From the order-preserving property of T_s we get $T_s\, a \preceq T_{t+s}\, \bar{x} \preceq T_s\, b$. Since $\preceq$ is closed, taking the limit with respect to t and applying Lemma A.2 yields

$$T_s\, a \preceq x^* \preceq T_s\, b \quad \text{for all } \ s \in \mathbb{T}. \tag{A.4}$$

In addition, from (A.4) and the order-preserving property of T_t,

$$T_{u+t}\, a \preceq T_t\, x^* \preceq T_{u+t}\, b \quad \text{for all } \ u, t \in \mathbb{T}. \tag{A.5}$$

Combining (A.4) and (A.5) with the diagonal property (2.1), we get

$$d(x^*, T_t\, x^*) \leq d(T_{u+t}\, a, T_{u+t}\, b) \quad \text{for all } \ u \in \mathbb{T}. \tag{A.6}$$

Taking the limit in u and using asymptotic contractivity yields $T_t\, x^* = x^*$. □

Proof of Theorem 2.1. We have already shown that x^* is stationary for (T_t) in Lemma A.3. Regarding convergence, fix $x \in \mathsf{X}$. Since (T_t) is asymptotically contractive and x^* is stationary, we have $d(T_t\, x, x^*) = d(T_t\, x, T_t\, x^*) \to 0$ as $t \to \infty$. This concludes the proof of Theorem 2.1. □

A.2. Proof of Theorem 2.2

We only used closedness of $\preceq$ in Lemma A.3, when we proved that the limit point x^* of the order-bounded trajectory $(T_t\, \bar{x})$ is a fixed point of each T_t. We now prove this without closedness by fixing $t \in \mathbb{T}$ and applying the triangle inequality to obtain

$$d(x^*, T_t\, x^*) \leq d(x^*, T_{u+t}\, \bar{x}) + d(T_{u+t}\, \bar{x}, T_t\, x^*).$$

The first term converges to zero in u by Lemma A.2. If we can show that the second term also converges to zero in u then we are done. This holds because T_t is continuous and $T_u\, \bar{x} \to x^*$, so, as $u \to \infty$,

$$d(T_{u+t}\, \bar{x}, T_t\, x^*) = d(T_t T_u\, \bar{x}, T_t\, x^*) \to d(T_t\, x^*, T_t\, x^*) = 0.$$

A.3. Proof of Proposition 3.1 and Theorem 3.1

Let the conditions of Theorem 3.1 hold. Because S is a partially ordered Polish space and Assumption 3.1 is in force, $(\mathscr{P}(S), \beta)$ is complete [16, Theorem 4.1] and $\preceq_{\mathrm{sd}}$ is closed [38]. We begin with the following lemma.

Lemma A.4. *If $\Lambda \subset \mathscr{P}(S)$ is tight, then there exists an increasing sequence of order intervals $(I_n)_{n \in \mathbb{Z}_+}$ such that $\sup_{\phi \in \Lambda} \phi(I_n^c) \to 0$ as $n \to \infty$.*

Proof. Take $(\epsilon_n) \subset (0, 1]$ with $\epsilon_n \downarrow 0$. By tightness, we can choose compact set K_0 with $\sup_{\phi \in \Lambda} \phi(K_0^c) \leq \epsilon_0$. By Assumption 3.1, compact sets are order bounded, so we can also take an order interval I_0 containing K_0 with $\sup_{\phi \in \Lambda} \phi(I_0^c) \leq \epsilon_0$. Next, we choose a compact set K_1 with $\sup_{\phi \in \Lambda} \phi(K_1^c) \leq \epsilon_1$, and then, using Assumption 3.1 again, an order interval I_1 such that $I_0 \cup K_1 \subset I_1$. This order interval obeys $I_0 \subset I_1$ and $\sup_{\phi \in \Lambda} \phi(I_1^c) \leq \epsilon_1$. Continuing in this way produces an increasing sequence of order intervals $(I_n)_{n \in \mathbb{Z}_+}$ with $\sup_{\phi \in \Lambda} \phi(I_n^c) \to 0$. □

Proof of Proposition 3.1. Suppose first that Λ is tight. We show that there exists a pair of distributions $\ell, u \in \mathscr{P}(S)$ such that $\ell \preceq_{\mathrm{sd}} \phi \preceq_{\mathrm{sd}} u$ for all $\phi \in \Lambda$. By Lemma A.4, we can take an increasing sequence of order intervals $(I_n)_{n \in \mathbb{Z}_+}$ and a sequence $(\epsilon_n) \subset (0, 1]$ with $\epsilon_n \downarrow 0$ and $\sup_{\phi \in \Lambda} \phi(I_n^c) \leq \epsilon_n$ for all $n \in \mathbb{Z}_+$. Let $I_n = [x_n, y_n]$. Since (I_n) is increasing, the sequence (x_n) is decreasing and (y_n) is increasing. For the distributions ℓ and u, we set

$$\ell := \sum_{n=0}^{\infty} \alpha_n \delta_{x_n} \quad \text{and} \quad u := \sum_{n=0}^{\infty} \alpha_n \delta_{y_n}, \quad \text{where} \quad \alpha_n := \epsilon_{n-1} - \epsilon_n.$$

In the definition above, $\epsilon_{-1} := 1$ and δ_z is the probability measure concentrated on z. One easily confirms that both ℓ and u are elements of $\mathscr{P}(S)$.

Now let J be any increasing set in $\mathscr{B}$. We claim that $\phi(J) \leq u(J)$ for all $\phi \in \Lambda$. To see that this is so, fix $\phi \in \Lambda$ and suppose first that $y_m \in J^c$ for all m. If this is so, then $\phi(J) = 0$. Indeed, fixing $m \in \mathbb{Z}_+$, the property $y_m \in J^c$ implies $J \subset I_m^c$. (Otherwise there is an $x \in J \cap I_m$ and, since J is increasing and $x \leq y_m$, we have $y_m \in J$.) This means that $\phi(J) \leq \phi(I_m^c) \leq \epsilon_m$. Taking the limit in m yields $\phi(J) = 0$. Hence $\phi(J) \leq u(J)$.

Now consider the other case, where $m := \min\{n \in \mathbb{Z}_+ \,:\, y_n \in J\}$ is finite. If $m = 0$, then $\phi(J) \leq 1 = u(J)$, so let $m \geq 1$. Observe that $J \cap I_{m-1}$ is empty. (If x is in both, then, since $x \leq y_{m-1}$ and $x \in J$, we have $y_{m-1} \in J$, which contradicts the definition of m.) As a result,

$$\phi(J) = \phi(J \cap I_{m-1}) + \phi(J \cap I_{m-1}^c) = \phi(J \cap I_{m-1}^c) \leq \phi(I_{m-1}^c) \leq \epsilon_{m-1}.$$

At the same time, since (y_n) is increasing, $u(J) = \epsilon_{m-1} - \epsilon_m + \epsilon_m - \epsilon_{m+1} + \epsilon_{m+1} - \epsilon_{m+2} + \cdots = \epsilon_{m-1}$. As a consequence, we have $\phi(J) \leq u(J)$. This shows that $\phi \preceq_{\mathrm{sd}} u$. A similar argument shows that $\ell \preceq_{\mathrm{sd}} \phi$ also holds. Hence Λ is order bounded.

Conversely, suppose Λ is order bounded, so there exist $\ell, u \in \mathscr{P}(S)$ with $\ell \preceq_{\mathrm{sd}} \phi \preceq_{\mathrm{sd}} u$ for all $\phi \in \Lambda$. Fix $\epsilon > 0$. Since ℓ and u are distributions on a Polish space, they are tight, so we can choose compact sets K_ℓ and K_u with $\ell(K_\ell^c) \leq \epsilon/2$ and $u(K_u^c) \leq \epsilon/2$. By Assumption 3.1, the compact set $K_\ell \cup K_u$ is order bounded, so there exist $a, b \in S$ with $K_\ell \cup K_u \subset [a, b]$. By Assumption 3.1 again, $[a, b]$ is compact.

Let $U_a := \{x \in S \,:\, x \geq a\}$ and $D_b := \{x \in S \,:\, x \leq b\}$. Then $[a, b] = U_a \cap D_b$ and $[a, b]^c = U_a^c \cup D_b^c$. Fix $\phi \in \Lambda$. Since U_a is increasing, we have $\ell(U_a) \leq \phi(U_a)$, giving $\phi(U_a^c) \leq \ell(U_a^c)$. Similarly, since D_b^c is increasing and $\phi \preceq_{\mathrm{sd}} u$, we have $\phi(D_b^c) \leq u(D_b^c)$. Now $K_\ell \subset [a, b]$ implies $U_a^c \subset [a, b]^c \subset K_\ell^c$, so $\ell(U_a^c) \leq \epsilon/2$. Likewise, $u(D_b^c) \leq \epsilon/2$. Therefore,

$$\phi([a, b]^c) \leq \phi(U_a^c) + \phi(D_b^c) \leq \ell(U_a^c) + u(D_b^c) \leq \epsilon.$$

Since $[a, b]$ is compact and this holds for all $\phi \in \Lambda$, we conclude that Λ is tight. □

Proof of Theorem 3.1. Let the conditions of Theorem 3.1 hold.

((i) ⇔ (ii)) This equivalence is immediate from Proposition 3.1.

((iii) $\implies$ (i)) If (T_t) is globally stable, then (T_t) is asymptotically contractive by Theorem 2.1. Moreover, when ψ^* is the unique stationary distribution, the constant trajectory $(T_t\psi^*) = (\psi^*)$ is trivially order bounded.

((i) $\implies$ (iii)) Suppose that (T_t) is asymptotically contractive and has an order-bounded trajectory. The space $(\mathscr{P}(S), \beta)$ is complete and β satisfies the diagonal property by Lemma 3.2. Moreover, $\preceq_{\mathrm{sd}}$ is closed with respect to β by Lemma 3.3. Hence (T_t) is globally stable by Theorem 2.1. □

A.4. Order-Theoretic Mixing and Asymptotic Contractivity

This section shows that weak order mixing implies asymptotic contractivity in the Bhattacharya metric. Proposition A.1 is the key step in the proofs of Theorem 5.1 and Theorem 5.2 (Appendix A.5), where it supplies the asymptotic contractivity needed to apply Theorem 3.1.

Proposition A.1. *Let P be an increasing stochastic kernel on S. If P is weakly order mixing, then P is asymptotically contractive on $\mathscr{P}(S)$ with respect to the metric β.*

Proof of Proposition A.1. Let $((X_k), (X_k'))$ and $(\mathcal{F}_k)$ be as in the definition of weak order mixing, and let $\tau := \inf\{k \in \mathbb{Z}_+ \,:\, X_k \leq X_k'\}$. Since $\leq$ is closed, $\{\tau = k\} \in \mathcal{F}_k$ for each k, and $\mathbb{P}_{x,x'}\{\tau < \infty\} = 1$ for all $x, x' \in S$. Fix $x, x' \in S$, $t \in \mathbb{Z}_+$ and $h \in ibS$ with $|h| \leq 1$. Since (X_k) and (X_k') are both P-Markov and defined on the same probability space, we have

$$\delta_x P^t h - \delta_{x'} P^t h = \mathbb{E}_{x,x'}[h(X_t) - h(X_t')] = A + B,$$

where

$$A := \sum_{k=0}^{t} \mathbb{E}_{x,x'}\left[\mathbb{1}_{\{\tau=k\}}(h(X_t) - h(X_t'))\right] \quad \text{and} \quad B := \mathbb{E}_{x,x'}\left[\mathbb{1}_{\{\tau>t\}}(h(X_t) - h(X_t'))\right].$$

Fix $k \leq t$. Since $\{\tau = k\} \in \mathcal{F}_k$ and both processes are P-Markov with respect to $(\mathcal{F}_k)$, conditioning on $\mathcal{F}_k$ gives

$$\mathbb{E}_{x,x'}\left[\mathbb{1}_{\{\tau=k\}}(h(X_t) - h(X_t'))\right] = \mathbb{E}_{x,x'}\left[\mathbb{1}_{\{\tau=k\}}\left((P^{t-k}h)(X_k) - (P^{t-k}h)(X_k')\right)\right].$$

Since $P^{t-k}h$ is increasing and $X_k \leq X'_k$ on $\{\tau = k\}$, this term is nonpositive. Hence $A \leq 0$. Moreover,

$$B \leq \mathbb{E}_{x,x'}\left[\mathbb{1}_{\{\tau>t\}}|h(X_t) - h(X'_t)|\right] \leq 2\mathbb{P}_{x,x'}\{\tau > t\}.$$

Hence $\delta_x P^t h - \delta_{x'} P^t h \leq 2\mathbb{P}_{x,x'}\{\tau > t\}$. Reversing the roles of x and x' gives $\delta_{x'} P^t h - \delta_x P^t h \leq 2\mathbb{P}_{x',x}\{\tau > t\}$, and hence

$$|\delta_x P^t h - \delta_{x'} P^t h| \leq 2 \max(\mathbb{P}_{x,x'}\{\tau > t\}, \mathbb{P}_{x',x}\{\tau > t\}).$$

Since $h \in ibS$ with $|h| \leq 1$ was arbitrary, we can take the supremum to get

$$\beta(\delta_x P^t, \delta_{x'} P^t) \leq 2 \max(\mathbb{P}_{x,x'}\{\tau > t\}, \mathbb{P}_{x',x}\{\tau > t\}) \quad \text{for all } x, x' \in S. \tag{A.7}$$

We wish to extend this to arbitrary (nondegenerate) initial conditions. To this end, fix $\phi, \psi \in \mathscr{P}(S)$. Let (X, X') be any coupling with marginals ϕ and ψ. For $h \in ibS$ with $|h| \leq 1$, we have

$$(\phi P^t)(h) - (\psi P^t)(h) = \mathbb{E}[(P^t h)(X)] - \mathbb{E}[(P^t h)(X')] = \mathbb{E}[(\delta_X P^t)(h) - (\delta_{X'} P^t)(h)].$$

Taking absolute values and using (A.7), we get

$$|(\phi P^t)(h) - (\psi P^t)(h)| \leq \mathbb{E}[\beta(\delta_X P^t, \delta_{X'} P^t)].$$

The right-hand side does not depend on h, so taking the supremum over $h \in ibS$ with $|h| \leq 1$ gives $\beta(\phi P^t, \psi P^t) \leq \mathbb{E}[\beta(\delta_X P^t, \delta_{X'} P^t)]$. By (A.7),

$$\beta(\phi P^t, \psi P^t) \leq 2\mathbb{E}[\max(\mathbb{P}_{X,X'}\{\tau > t\}, \mathbb{P}_{X',X}\{\tau > t\})]. \tag{A.8}$$

Since $\mathbb{P}_{x,x'}\{\tau < \infty\} = 1$ for all $x, x' \in S$, both $\mathbb{P}_{x,x'}\{\tau > t\}$ and $\mathbb{P}_{x',x}\{\tau > t\}$ converge to 0 as $t \to \infty$ for each (x, x'). By the dominated convergence theorem, the right-hand side of (A.8) converges to 0. Since $\phi, \psi \in \mathscr{P}(S)$ were arbitrary, P is asymptotically contractive. □

A.5. Proof of Theorem 5.1 and Theorem 5.2

We begin with a lemma showing that convergence in the Bhattacharya metric implies tightness.

Lemma A.5. *If $(\phi_n)_{n\in\mathbb{Z}_+} \subset \mathscr{P}(S)$ and $\beta(\phi_n, \phi) \to 0$ for some $\phi \in \mathscr{P}(S)$, then $\{\phi_n\}_{n\in\mathbb{Z}_+}$ is tight.*

Proof. Fix $\epsilon > 0$. Since ϕ is a distribution on a Polish space, it is tight, so we can choose a compact set K with $\phi(K^c) \leq \epsilon$. By Lemma 3.1, there exist $a, b \in S$ with $K \subset [a, b]$, and $[a, b]$ is compact. As in the proof of Proposition 3.1, let $U_a := \{x \in S : x \geq a\}$ and $D_b := \{x \in S : x \leq b\}$. Both sets are closed, since $\leq$ is closed. Moreover, D_b^c is increasing and U_a^c is decreasing, so $\mathbb{1}_{D_b^c}$ and $-\mathbb{1}_{U_a^c}$ are elements of ibS bounded in absolute value by one. Convergence in β therefore gives

$$\phi_n(D_b^c) \to \phi(D_b^c) \leq \phi(K^c) \leq \epsilon \quad \text{and} \quad \phi_n(U_a^c) \to \phi(U_a^c) \leq \phi(K^c) \leq \epsilon.$$

Since $[a, b]^c = U_a^c \cup D_b^c$, we can choose $N \in \mathbb{Z}_+$ such that $\phi_n([a, b]^c) \leq 3\epsilon$ for all $n \geq N$. Each of the finitely many distributions $\phi_0, \ldots, \phi_{N-1}$ is tight, so there exist compact sets $K_0, \ldots, K_{N-1}$ with $\phi_n(K_n^c) \leq \epsilon$ for $n < N$. The set $\bar{K} := [a, b] \cup K_0 \cup \cdots \cup K_{N-1}$ is compact and satisfies $\phi_n(\bar{K}^c) \leq 3\epsilon$ for all $n \in \mathbb{Z}_+$. Hence $\{\phi_n\}$ is tight. □

Next we prove the discrete time result in Theorem 5.1.

Proof of Theorem 5.1. Let P be any stochastic kernel on S. If P is globally stable on $\mathscr{P}(S)$, then P is bounded in probability. Indeed, for each $x \in S$, the sequence of distributions $t \mapsto \delta_x P^t = P^t(x, \cdot)$ converges in β to the unique stationary distribution, and hence is tight by Lemma A.5.

Now suppose that P is increasing, order reversing, and bounded in probability. By Lemma A.5 in the appendix of [13], there exists an $(S \times S)$-valued Markov process $((X_t), (X'_t))_{t\in\mathbb{Z}_+}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $(X_t)_{t\in\mathbb{Z}_+}$ and $(X'_t)_{t\in\mathbb{Z}_+}$ are independent, both of these processes are P-Markov, and $\tau := \inf\{t \geq 0 : X_t \leq X'_t\}$ obeys $\mathbb{P}_{x,x'}\{\tau < \infty\} = 1$ for all $x, x' \in S$. By independence, (X_t) and (X'_t) are both P-Markov with respect to the natural filtration of $((X_t), (X'_t))$, so P is weakly order mixing. Proposition A.1 now implies that P is asymptotically contractive on $\mathscr{P}(S)$ with respect to the metric β. Global stability now follows from Theorem 3.1 (ii), applied with $\mathbb{T} = \mathbb{Z}_+$, given that boundedness in probability implies existence of a tight trajectory. □

Now we prove the continuous time result in Theorem 5.2.

Proof of Theorem 5.2. Since (P_t) is weakly order mixing, there exists a $u > 0$ such that P_u is weakly order mixing. Since P_u is also increasing, Proposition A.1 implies that P_u is asymptotically contractive on $(\mathscr{P}(S), \beta)$. To extend this to (P_t), fix $\phi, \psi \in \mathscr{P}(S)$ and $t \geq 0$, and write $t = nu + r$ with $n \in \mathbb{Z}_+$ and $r \in [0, u)$. Since $P_t = P_{nu} P_r$, $P_{nu} = (P_u)^n$ and P_r is increasing, Lemma 4.1 gives

$$\beta(\phi P_t, \psi P_t) = \beta((\phi P_u^n) P_r, (\psi P_u^n) P_r) \leq \beta(\phi P_u^n, \psi P_u^n).$$

As $t \to \infty$ we have $n \to \infty$, so the right-hand side converges to zero. Hence (P_t) is asymptotically contractive on $(\mathscr{P}(S), \beta)$. If (P_t) is globally stable with unique stationary distribution ψ^*, then the trajectory $t \mapsto \psi^* P_t = \psi^*$ is constant, and hence trivially tight. Conversely, if (P_t) has at least one tight trajectory, then global stability follows from Theorem 3.1, since (P_t) is order-preserving and asymptotically contractive. □

A.6. Proof of Theorem 6.1

We prove this theorem using a sequence of lemmas.

Lemma A.6. *If Assumption 6.1 holds, then (P_t) is increasing.*

Proof. Fix $x, x' \in S$ with $x \leq x'$. On a common probability space, construct two PDMPs (X_t) and (X_t') starting from x and x' respectively, using the same jump times (T_n) and the same shocks (ξ_n). The embedded chains satisfy $Z_0 = x \leq x' = Z_0'$. By induction: if $Z_n \leq Z_n'$, then $\Phi(Z_n, E_{n+1}) \leq \Phi(Z_n', E_{n+1})$ by Assumption 6.1 (i) and

$$Z_{n+1} = F(\Phi(Z_n, E_{n+1}), \xi_{n+1}) \leq F(\Phi(Z_n', E_{n+1}), \xi_{n+1}) = Z_{n+1}'$$

by Assumption 6.1 (ii). Hence $Z_n \leq Z_n'$ for all n. For any $t \geq 0$, applying Assumption 6.1 (i) gives

$$X_t = \Phi(Z_{N_t}, t - T_{N_t}) \leq \Phi(Z_{N_t}', t - T_{N_t}) = X_t'.$$

Therefore, for any $h \in ibS$, $P_t h(x) = \mathbb{E}_x h(X_t) \leq \mathbb{E}_{x'} h(X_t') = P_t h(x')$. □

Lemma A.7. *If Assumptions 6.1 and 6.2 hold, then (P_t) is bounded in probability.*

Proof. Fix $x_0 \in S$ and $\varepsilon > 0$. We need to find a compact set D such that $\mathbb{P}_{x_0}\{X_t \in D\} \geq 1 - \varepsilon$ for all $t \geq 0$. By Assumption 6.2, $f_1(\xi_n) \leq Z_n \leq f_2(\xi_n)$ for all $n \geq 1$, so each post-jump state lies in $[f_1(\xi_n), f_2(\xi_n)]$. Choose a compact $K \subset S$ with $\mathbb{P}\{f_i(\xi) \in K\} \geq 1 - \varepsilon/6$ for $i = 1, 2$. By Assumption 3.1, $K \subset [a, b]$ for some $a, b \in S$, and $C \coloneqq [a, b]$ is compact. Whenever both $f_1(\xi_n)$ and $f_2(\xi_n)$ fall in K, the post-jump state Z_n lies in $K \subset C$.

Next, we show that Z_{N_t}, the state at the most recent jump before time t, lies in C with high probability. Formally, ξ_n is independent of $(E_1, \dots, E_{n+1})$ and $\{N_t = n\}$ is determined by $(E_1, \dots, E_{n+1})$, so

$$\mathbb{P}_{x_0}\{Z_{N_t} \notin C,\ N_t \geq 1\} = \sum_{n=1}^{\infty} \mathbb{P}\{N_t = n,\ Z_n \notin C\} \leq \frac{\varepsilon}{3} \sum_{n=1}^{\infty} \mathbb{P}\{N_t = n\} = \frac{\varepsilon}{3}.$$

Between consecutive jumps, the process follows the flow Φ, so $X_t = \Phi(Z_{N_t}, R_t)$ where $R_t \coloneqq t - T_{N_t}$ is the time elapsed since the last jump. The event $\{R_t > M\}$ requires no Poisson arrivals in the interval $(t - M, t]$, so $\mathbb{P}\{R_t > M\} \leq e^{-\lambda M}$. Choose $M > 0$ with $e^{-\lambda M} < \varepsilon/3$. Before the first jump ($N_t = 0$), the process follows its initial trajectory: $X_t = \Phi(x_0, t)$. The probability of no jump by time t is $\mathbb{P}\{N_t = 0\} = e^{-\lambda t}$, which is small for large t. Choose $T > 0$ with $e^{-\lambda T} < \varepsilon/3$.

Since Φ is continuous,

$$D \coloneqq \Phi(C \times [0, M]) \cup \Phi(\{x_0\} \times [0, T])$$

is compact. (The first component captures the flow starting from any post-jump state in C and running for at most M time units. The second captures the initial trajectory from x_0 up to time T, covering the possibility that no jump has yet occurred.) Consider first the case $t \leq T$: if $N_t = 0$, then $X_t = \Phi(x_0, t) \in D$. For any such t: on the event

$\{N_t \geq 1,\ Z_{N_t} \in C,\ R_t \leq M\}$, we have $X_t = \Phi(Z_{N_t}, R_t) \in D$. The three complementary events each have small probability:

$$\mathbb{P}_{x_0}\{X_t \notin D\} \leq \mathbb{P}\{N_t = 0\} + \mathbb{P}\{Z_{N_t} \notin C,\ N_t \geq 1\} + \mathbb{P}\{R_t > M\} \leq e^{-\lambda t} + \frac{\varepsilon}{3} + e^{-\lambda M}.$$

For $t > T$ each term is at most $\varepsilon/3$, giving a total bound of ε. For $t \leq T$ the first term covers the event $\{N_t = 0\}$ (on which $X_t \in D$ by construction), so $\mathbb{P}_{x_0}\{X_t \notin D\} \leq \varepsilon/3 + \varepsilon/3 < \varepsilon$. ☐

Lemma A.8. *If Assumptions 6.1 and 6.2 hold, then (P_t) is weakly order mixing.*

Proof. Fix $u > 0$. We show that P_u is weakly order mixing. Let (E_n) be IID $\text{Exp}(\lambda)$, and let (ξ_n) and (ξ'_n) be IID with common distribution μ, all mutually independent. Fix $x, x' \in S$ and define embedded chains by $Z_0 = x$, $Z'_0 = x'$ and, for $n \geq 1$,

$$Z_n = F(\Phi(Z_{n-1}, E_n), \xi_n), \qquad Z'_n = F(\Phi(Z'_{n-1}, E_n), \eta_n),$$

where $\eta_n := \xi_n$ if $Z_k \leq Z'_k$ for some $k < n$ and $\eta_n := \xi'_n$ otherwise. Thus, the two chains use independent shocks until they first become ordered and common shocks afterwards. Since the choice between ξ_n and ξ'_n depends only on $(E_k, \xi_k, \xi'_k)_{k<n}$, the sequence (η_n) is IID with distribution μ and independent of (E_n). Let (X_t) and (X'_t) be the PDMPs with common jump times $T_n := E_1 + \cdots + E_n$ and embedded chains (Z_n) and (Z'_n).

The process $((X_t), (X'_t), (I_t))$, where I_t is the indicator that $Z_k \leq Z'_k$ for some k with $T_k \leq t$, is itself a PDMP on $S \times S \times \{0, 1\}$, and hence a Markov process [40, Theorem 2.4]. Whatever the value of I_t, the future of each of (X_t) and (X'_t) is that of a PDMP with characteristics (Φ, λ, μ, F) started at its current value. Hence, if $(\mathcal{F}_k)$ is the natural filtration of $((X_{ku}), (X'_{ku}), (I_{ku}))_{k \in \mathbb{Z}_+}$, then (X_{ku}) and (X'_{ku}) are both P_u-Markov with respect to $(\mathcal{F}_k)$.

Let $N := \inf\{n \geq 0 : Z_n \leq Z'_n\}$. Since weak order mixing is defined through the skeleton, we need the ordering to be visible at a time of the form ku; for this reason the chains switch to common shocks once ordered, so that the ordering persists. On $\{N < \infty\}$, the two chains use common shocks after step N, so the induction in the proof of Lemma A.6 gives $X_t \leq X'_t$ for all $t \geq T_N$. In particular, $X_{ku} \leq X'_{ku}$ for all k with $ku \geq T_N$. It remains to show that $N < \infty$ with probability one.

Let f_1 and f_2 be as in Assumption 6.2 and set $\delta := \mathbb{P}\{f_2(\xi) \leq f_1(\xi')\} > 0$. If $N > n - 1$, then $\eta_n = \xi'_n$, and on the event $\{f_2(\xi_n) \leq f_1(\xi'_n)\}$ we have $Z_n = F(\Phi(Z_{n-1}, E_n), \xi_n) \leq f_2(\xi_n) \leq f_1(\xi'_n) \leq F(\Phi(Z'_{n-1}, E_n), \xi'_n) = Z'_n$. Hence $\{N > k\} \subset \bigcap_{n=1}^k \{f_2(\xi_n) > f_1(\xi'_n)\}$. Since the pairs (ξ_n, ξ'_n) are IID, it follows that $\mathbb{P}\{N > k\} \leq (1 - \delta)^k$, and therefore $N < \infty$ almost surely. ☐

Proof of Theorem 6.1. By Lemma A.6, (P_t) is increasing. By Lemma A.7, (P_t) is bounded in probability, which implies the existence of a tight trajectory. By Lemma A.8, (P_t) is weakly order mixing. Global stability now follows from Theorem 5.2. ☐

A.7. Proof of Theorem 5.3

Let the conditions of Theorem 5.3 hold. We write $\lfloor \cdot \rfloor$ for the floor function.

Lemma A.9. *If the conditions of Theorem 5.3 hold, then*

$$(1 - \epsilon)\delta_a + \epsilon \delta_{\hat{x}} \preceq_{sd} P_u(a, \cdot) \quad \textit{and} \quad P_u(b, \cdot) \preceq_{sd} (1 - \epsilon)\delta_b + \epsilon \delta_{\hat{x}} \tag{A.9}$$

Proof. See the proof of Eq. (2) in Theorem 2 of [11]. ☐

Lemma A.10. *If the conditions of Theorem 5.3 hold, then*

$$\beta(\phi P_u^n, \psi P_u^n) \leq 2(1 - \epsilon)^n \quad \textit{for all } \phi, \psi \in \mathscr{P}(S) \textit{ and } n \in \mathbb{Z}_+. \tag{A.10}$$

Proof. We set $P = P_u$ for brevity. Fix $h \in ibS$ with $|h| \leq 1$. Define $d_n = P^n h(a)$ and $e_n = P^n h(b)$. Since P is increasing and h is increasing, $P^{n-1} h$ is also increasing for all $n \geq 1$. Applying the bounds from Lemma A.9 to the increasing function $P^{n-1} h$ yields

$$d_n = (P(P^{n-1} h))(a) \geq (1 - \epsilon) d_{n-1} + \epsilon (P^{n-1} h)(\hat{x})$$

$$e_n = (P(P^{n-1}h))(b) \leq (1-\epsilon)e_{n-1} + \epsilon(P^{n-1}h)(\hat{x})$$

Subtracting one inequality from the other, we get $e_n - d_n \leq (1-\epsilon)(e_{n-1} - d_{n-1})$, and hence

$$e_n - d_n \leq (1-\epsilon)^n(e_0 - d_0) = (1-\epsilon)^n(h(b) - h(a)) \leq 2(1-\epsilon)^n. \tag{A.11}$$

Since a is the bottom and b is the top of S, we have $(P^n h)(a) \leq (P^n h)(x) \leq (P^n h)(b)$ for all $x \in S$, so, integrating with respect to $\phi \in \mathscr{P}(S)$, we have $(P^n h)(a) \leq \phi(P^n h) \leq (P^n h)(b)$. Therefore, for any two probability measures ϕ and ψ,

$$|\phi(P^n h) - \psi(P^n h)| \leq (P^n h)(b) - (P^n h)(a) \leq 2(1-\epsilon)^n.$$

Taking the supremum over all $h \in ibS$ with $|h| \leq 1$ proves the claim in the lemma. $\square$

Proof of Theorem 5.3. Let the stated conditions hold. We first claim that $(P_t)_{t\in\mathbb{T}}$ is asymptotically contractive. For $\mathbb{T} = \mathbb{R}_+$, fix $t \in (0,\infty)$, write $t = nu + r$ where $n = \lfloor t/u \rfloor$ and $0 \leq r < u$. By the semigroup property, we have $P_t = P_{nu+r} = P_{nu}P_r = P_u^n P_r$, so

$$\beta(\phi P_t, \psi P_t) = \beta(\phi P_u^n P_r, \psi P_u^n P_r) \leq \beta(\phi P_u^n, \psi P_u^n),$$

where the inequality is by the nonexpansiveness in Lemma 4.1. Applying the contraction bound in Lemma A.10, we obtain

$$\beta(\phi P_u^n, \psi P_u^n) \leq 2(1-\epsilon)^n = 2(1-\epsilon)^{\lfloor t/u \rfloor}.$$

As $t \to \infty$, $\lfloor t/u \rfloor \to \infty$, so $(P_t)_{t\geq 0}$ is asymptotically contractive, as claimed.

For $\mathbb{T} = \mathbb{Z}_+$, the proof is similar. We write $t = nu + r$ where $n = \lfloor t/u \rfloor$ and r is an integer obeying $0 \leq r < u$. Since $P_t = P_u^n P_r$, we can again use Lemmas 4.1 and A.10 to obtain

$$\beta(\phi P_t, \psi P_t) \leq \beta(\phi P_u^n, \psi P_u^n) \leq 2(1-\epsilon)^n \to 0$$

as $t \to \infty$. Hence $(P_t)_{t\in\mathbb{Z}_+}$ is asymptotically contractive.

To complete the proof of global stability, we use Theorem 3.1. For any increasing $h \in ibS$ and any $\phi \in \mathscr{P}(S)$, we have $h(a) \leq \phi(h) \leq h(b)$, and hence $\delta_a \preceq_{\text{sd}} \phi \preceq_{\text{sd}} \delta_b$. In particular, for any $\phi \in \mathscr{P}(S)$ and any $t \in \mathbb{T}$, we have $\delta_a \preceq_{\text{sd}} \phi P_t \preceq_{\text{sd}} \delta_b$. Thus every trajectory of (P_t) is order bounded. Since (P_t) is asymptotically contractive, increasing, and has an order-bounded trajectory, Theorem 3.1 implies that (P_t) is globally stable.

For the exponential bound (5.2), let ϕ^* be the unique stationary distribution. Since $\lfloor t/u \rfloor \geq t/u - 1$, we have $(1-\epsilon)^{\lfloor t/u \rfloor} \leq (1-\epsilon)^{t/u-1}$. Writing $(1-\epsilon)^{t/u} = e^{-\alpha t}$ where $\alpha = \ln(1/(1-\epsilon))/u > 0$, we obtain

$$\beta(\phi P_t, \psi P_t) \leq \frac{2}{1-\epsilon} e^{-\alpha t}.$$

Setting $\psi = \phi^*$ and using $\phi^* P_t = \phi^*$ yields the bound in (5.2).

The monotone ergodicity result in Theorem 5.3 follows from Theorem 3.1 and Proposition 4.1 of [14]. $\square$

References

[1] L. E. Dubins, D. A. Freedman, Invariant probabilities for certain Markov processes, The Annals of Mathematical Statistics 37 (4) (1966) 837–848.

[2] T. Kamae, U. Krengel, G. O'Brien, Stochastic inequalities on partially ordered spaces, The Annals of Probability (1977) 899–912.

[3] T. Lindvall, Lectures on the coupling method, Courier Corporation, 2002.

[4] R. Bhattacharya, M. Majumdar, N. Hashimzade, Limit theorems for monotone Markov processes, Sankhya A 72 (1) (2010) 170–190.

[5] V. Strassen, The existence of probability measures with given marginals, The Annals of Mathematical Statistics (1965) 423–439.

[6] J. G. Propp, D. B. Wilson, Exact sampling with coupled Markov chains and applications to statistical mechanics, Random Structures & Algorithms 9 (1-2) (1996) 223–252.

[7] J. A. Fill, M. Machida, Stochastic monotonicity and realizable monotonicity, The Annals of Probability 29 (2) (2001) 938–978.

[8] S. Asmussen, On coupling and weak convergence to stationarity, The Annals of Applied Probability 2 (3) (1992) 739–751.

[9] H. Thorisson, Coupling, Stationarity, and Regeneration, Springer, 2000.

[10] R. Bhattacharya, O. Lee, Asymptotics of a class of Markov processes which are not in general irreducible, The Annals of Probability (1988) 1333–1347.

[11] H. A. Hopenhayn, E. C. Prescott, Stochastic monotonicity and stationary distributions for dynamic economies, Econometrica (1992) 1387–1406.

[12] T. Kamihigashi, J. Stachurski, An order-theoretic mixing condition for monotone Markov chains, Statistics and Probability Letters 82 (2012) 262–267.

[13] T. Kamihigashi, J. Stachurski, Stochastic stability in monotone economies, Theoretical Economics 9 (2) (2014) 383–407. doi:10.3982/TE1367.

[14] T. Kamihigashi, J. Stachurski, Seeking ergodicity in dynamic economies, Journal of Economic Theory 163 (2016) 900–924. doi:10.1016/j.jet.2016.03.006.

[15] S. Foss, V. Shneer, J. P. Thomas, T. Worrall, Stochastic stability of monotone economies in regenerative environments, Journal of Economic Theory 173 (2018) 334–360. doi:10.1016/j.jet.2017.12.007.

[16] T. Kamihigashi, J. Stachurski, A unified stability theory for classical and monotone Markov chains, Journal of Applied Probability 56 (1) (2019) 1–22.

[17] S. Foss, M. Scheutzow, Compressibility and stochastic stability of monotone markov chain (2024). arXiv:2403.15259.
URL https://arxiv.org/abs/2403.15259

[18] B. Light, A note on the stability of monotone Markov chains, Operations Research Letters 57 (2024).

[19] B. Light, Invariant distributions in nonlinear Markov chains with aggregators: Theory, computation, and applications, Operations Research in press (2026).

[20] F. Daly, On Stein's method for stochastically monotone single-birth chains, Statistics & Probability Letters 206 (2024) 109993.

[21] M. Hairer, J. C. Mattingly, M. Scheutzow, Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations, Probability Theory and Related Fields 149 (1–2) (2011) 223–259.

[22] O. Butkovsky, M. Scheutzow, Couplings via comparison principle and exponential ergodicity of SPDEs in the hypoelliptic setting, Communications in Mathematical Physics 379 (3) (2020) 1001–1034.

[23] S. P. Meyn, R. L. Tweedie, Markov chains and stochastic stability, Cambridge University Press, 2009.

[24] J. Hahn, G. Kuersteiner, M. Mazzocco, Central limit theory for combined cross section and time series with an application to aggregate productivity shocks, Econometric Theory 40 (1) (2024) 162–212.

[25] Q. Ma, A. A. Toda, Optimal savings with preference for wealth, arXiv preprint arXiv:2509.12195 (2025).

[26] A. C. Ran, M. C. Reurings, A fixed point theorem in partially ordered sets and some applications to matrix equations, Proceedings of the American Mathematical Society (2004) 1435–1443.

[27] J. J. Nieto, R. L. Pouso, R. Rodríguez-López, Fixed point theorems in ordered abstract spaces, Proceedings of the American Mathematical Society (2007) 2505–2517.

[28] J. Caballero, J. Harjani, K. Sadarangani, Contractive-like mapping principles in ordered metric spaces and application to ordinary differential equations, Fixed Point Theory and Applications 2010 (2010) 1–14.

[29] H. Ben-El-Mechaiekh, The Ran–Reurings fixed point theorem without partial order: a simple proof, Journal of Fixed Point Theory and Applications 16 (2014) 373–383.

[30] J. Górnicki, Fixed point theorems in preordered sets, Journal of Fixed Point Theory and Applications 23 (4) (2021) 71.

[31] A. Orlik, L. Veldkamp, Understanding uncertainty shocks and the role of black swans, NBER Working Paper No. 20445 (2014).

[32] T. Cogley, T. J. Sargent, Drifting risk-adjusted inflation targets, Journal of Monetary Economics 55 (5) (2008) 740–757.

[33] D. Denisov, D. Korshunov, V. Wachtel, Potential analysis for positive recurrent Markov chains with asymptotically zero drift: Power-type asymptotics, Stochastic Processes and their Applications 123 (8) (2013) 3027–3051. doi:10.1016/j.spa.2013.04.011.

[34] S. Foss, D. Korshunov, S. Zachary, An Introduction to Heavy-Tailed and Subexponential Distributions, 2nd Edition, Springer Series in Operations Research and Financial Engineering, Springer, New York, 2013. doi:10.1007/978-1-4614-7101-1.

[35] X. Gabaix, J.-M. Lasry, P.-L. Lions, B. Moll, The dynamics of inequality, Econometrica 84 (6) (2016) 2071–2111.

[36] B. K. Beare, A. A. Toda, Determination of Pareto exponents in economic models driven by Markov multiplicative processes, Econometrica 90 (4) (2022) 1811–1833.

[37] A. C. Zaanen, Introduction to operator theory in Riesz spaces, Springer, 2012.

[38] T. Kamae, U. Krengel, Stochastic partial ordering, The Annals of Probability 6 (6) (1978) 1044–1049. doi:10.1214/aop/1176995392.

[39] A. Lasota, M. C. Mackey, Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics, 2nd Edition, Vol. 97 of Applied Mathematical Sciences, Springer-Verlag, New York, 1994.

[40] R. Rudnicki, M. Tyran-Kamińska, Piecewise deterministic processes in biological models, Vol. 1, Springer, 2017.

[41] N. L. Stokey, R. E. Lucas, Recursive methods in dynamic economics, Harvard University Press, 1989.

[42] D. Czapla, K. Horbacz, H. Wojewódka-Ściążko, Ergodic properties of some piecewise-deterministic Markov process with application to gene expression modelling, Stochastic Processes and their Applications 130 (5) (2020) 2851–2885.

[43] D. Czapla, On the existence and uniqueness of stationary distributions for some piecewise deterministic Markov processes with state-dependent jump intensity, Results in Mathematics 79 (5) (2024) 177.

[44] J. Suda, Belief shocks and the macroeconomy, banque de France Working Paper (2010).

[45] T. Kamihigashi, J. Stachurski, A necessary and sufficient condition for global stability of monotone Markov processes, arXiv preprint arXiv:2604.03979 (2026).

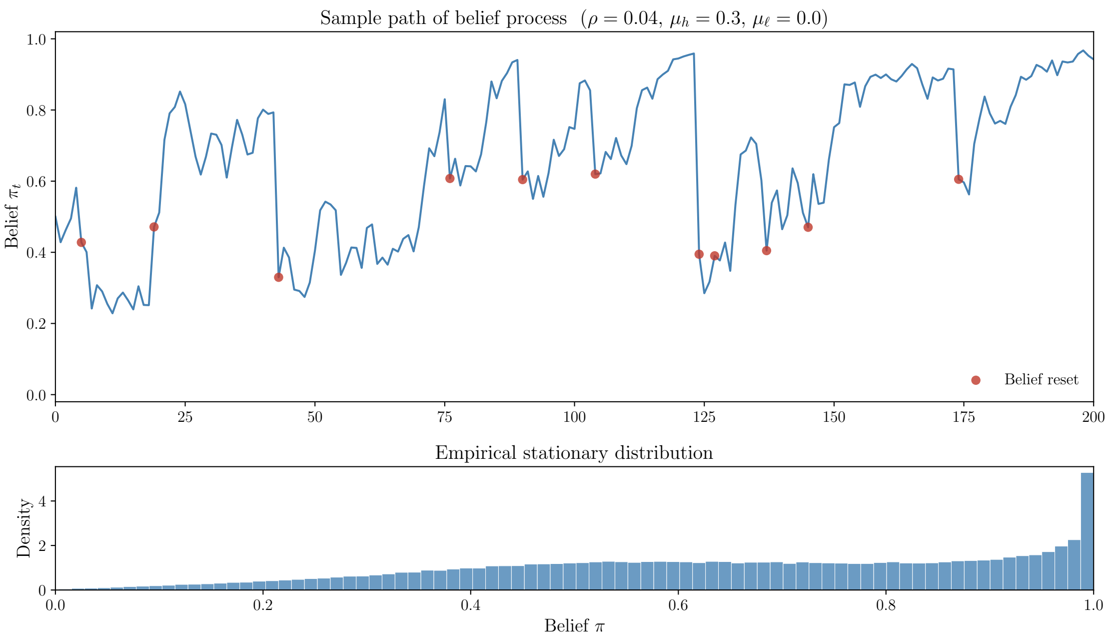

Figure 2: Simulation of the belief shock model in Section 7.1.