Introduction	Establishing drift	Using drift	Applications	References

# State-dependent Foster-Lyapunov criteria

Stephen Connor stephen.connor@york.ac.uk

Joint work with Gersende Fort, CNRS-TELECOM ParisTech; supported by CRiSM and the French National Research Agency

THE UNIVERSITY of fork

March 2012

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References
Outline				



Introduction

- The general problem
- Drift conditions
- Establishing a subsampled drift condition
   Examples
- Using a subsampled drift conditionInterplay of subsampling rates and moments

## Applications



Let  $\Phi = \{\Phi_n, n \ge 0\}$  be a time-homogeneous Markov chain on a state space X. (Assume that  $\Phi$  is phi-irreducible and aperiodic, for simplicity.)

We're interested in what can be said about the long-term behaviour of  $\Phi$ . For example:

- does  $\Phi$  converge to an equilibrium distribution?
- if so, in what norm does this convergence take place, and how fast?
- and how is this related to the average time spent between successive visits to certain sets?

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References
Notation				

- $P^n$ : *n*-step transition kernel of  $\Phi$ ;
- for a non-negative function f, and measure  $\mu$ :

$$P^n f(x) = \mathbb{E}_x[f(\Phi_n)], \qquad \mu(f) = \int f(y)\mu(dy);$$

$$\|\mu\|_{g} = \sup_{f:|f| \le g} |\mu(f)|, \qquad \|\cdot\|_{\mathrm{TV}} \equiv \|\cdot\|_{1};$$

$$P(x, \cdot) \ge \varepsilon \nu(\cdot)$$
 for all  $x \in C$ .

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References
Notation				

- $P^n$ : *n*-step transition kernel of  $\Phi$ ;
- for a non-negative function f, and measure  $\mu$ :

$$\mathcal{P}^n f(x) = \mathbb{E}_x[f(\Phi_n)], \qquad \mu(f) = \int f(y)\mu(dy);$$

$$\|\mu\|_{g} = \sup_{f:|f| \le g} |\mu(f)|, \qquad \|\cdot\|_{\mathrm{TV}} \equiv \|\cdot\|_{1};$$

$$P(x, \cdot) \ge \varepsilon \nu(\cdot)$$
 for all  $x \in C$ .

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References
Notation				

- $P^n$ : *n*-step transition kernel of  $\Phi$ ;
- for a non-negative function f, and measure  $\mu$ :

$$P^n f(x) = \mathbb{E}_x[f(\Phi_n)], \qquad \mu(f) = \int f(y)\mu(dy);$$

$$\|\mu\|_{g} = \sup_{f:|f| \le g} |\mu(f)|, \qquad \|\cdot\|_{\mathrm{TV}} \equiv \|\cdot\|_{1};$$

$$P(x, \cdot) \ge \varepsilon \nu(\cdot)$$
 for all  $x \in C$ .

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References
Notation				

- $P^n$ : *n*-step transition kernel of  $\Phi$ ;
- for a non-negative function f, and measure  $\mu$ :

$$\mathcal{P}^n f(x) = \mathbb{E}_x[f(\Phi_n)], \qquad \mu(f) = \int f(y)\mu(dy);$$

$$\|\mu\|_{g} = \sup_{f:|f| \le g} |\mu(f)|, \qquad \|\cdot\|_{\mathrm{TV}} \equiv \|\cdot\|_{1};$$

$$P(x, \cdot) \ge \varepsilon \nu(\cdot)$$
 for all  $x \in C$ .

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References
Notation				

- $P^n$ : *n*-step transition kernel of  $\Phi$ ;
- for a non-negative function f, and measure  $\mu$ :

$$P^n f(x) = \mathbb{E}_x[f(\Phi_n)], \qquad \mu(f) = \int f(y)\mu(dy);$$

$$\|\mu\|_{g} = \sup_{f:|f| \le g} |\mu(f)|, \qquad \|\cdot\|_{\mathrm{TV}} \equiv \|\cdot\|_{1};$$

• first return time to a set:  $\tau_{\mathcal{A}} = \inf\{n \ge 1 : \Phi_n \in \mathcal{A}\};$ •  $\mathcal{C}$  is a *small set* if  $\exists \varepsilon > 0$  and a measure  $\nu$  s.t.

$$P(x, \cdot) \ge \varepsilon \nu(\cdot)$$
 for all  $x \in C$ .



 $\Phi$  is called *ergodic* if it has a finite invariant measure  $\pi$  ( $\pi = \pi P$ ). In this case, for any x,

$$\|\mathcal{P}^n(x,\cdot)-\pi\|_{\mathrm{TV}} o 0$$
 as  $n o\infty.$ 

(1)

Equivalently, we can find a small set  $\ensuremath{\mathcal{C}}$  with

$$\sup_{x\in\mathcal{C}}\mathbb{E}_x[\tau_{\mathcal{C}}]<\infty.$$



 $\Phi$  is called *ergodic* if it has a finite invariant measure  $\pi$  ( $\pi = \pi P$ ). In this case, for any x,

$$\|P^n(x,\cdot)-\pi\|_{\mathrm{TV}} o 0$$
 as  $n o\infty.$ 

(1)

Equivalently, we can find a small set  $\ensuremath{\mathcal{C}}$  with

$$\sup_{x\in\mathcal{C}}\mathbb{E}_x[\tau_{\mathcal{C}}]<\infty.$$

But *how fast* does the convergence in (1) take place?

Introduction	Establishing drift	Using drift	Applications	References
000000	0	000		

## Definition

 $\Phi$  is *geometrically ergodic* if there exists r > 1 with

 $\left\|P^{n}(x,\cdot)-\pi(\cdot)\right\|_{\mathrm{TV}}\leq M_{x}r^{-n}.$ 

Introduction	Establishing drift	Using drift	Applications	References
000000	0	000		

#### Definition

 $\Phi$  is *geometrically ergodic* if there exists r > 1 with

$$\left\| \mathsf{P}^n(x,\cdot) - \pi(\cdot) \right\|_{\mathrm{TV}} \leq M_x r^{-n}.$$

## Equivalently:

• there exists a scale function  $V : X \to [1, \infty)$ , a small set C, and constants  $\beta \in (0, 1)$ ,  $b < \infty$ , with

$$\mathbb{E}_{x}\left[V(\Phi_{1})\right] = PV(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x);$$

• 
$$\sup_{x\in\mathcal{C}} \mathbb{E}_x[\beta^{-\tau_{\mathcal{C}}}] < \infty.$$



The inequality

$$PV(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x);$$

- is called a Foster-Lyapunov drift condition.
  - often easiest way of showing that  $\Phi$  is geometrically ergodic;
  - if V is bounded then  $\Phi$  is *uniformly ergodic*.



The inequality

$$PV(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x);$$

- is called a Foster-Lyapunov drift condition.
  - often easiest way of showing that  $\Phi$  is geometrically ergodic;
  - if V is bounded then  $\Phi$  is *uniformly ergodic*.

*Subgeometric* ergodicity is implied by a weaker drift condition:

$$PV(x) \leq V(x) - \phi \circ V(x) + b\mathbf{1}_{\mathcal{C}}(x)$$

for some concave non-negative function  $\phi$ .

Introduction 00000●	Establishing drift 0	Using drift 000	Applications	References
Questions:				



**(**) When can we find a function  $n: X \to \mathbb{N}$  such that

$$P^{n(x)}V(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x)$$
?

(2)



**(**) When can we find a function  $n : X \to \mathbb{N}$  such that

$$P^{n(x)}V(x) \le \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x)?$$
(2)

Alternatively, if (2) holds for some n and V, what can be said about moments of the return time to C?



**(**) When can we find a function  $n : X \to \mathbb{N}$  such that

$$P^{n(x)}V(x) \le \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x)?$$
(2)

- Alternatively, if (2) holds for some n and V, what can be said about moments of the return time to C?
- And when is this useful?!

#### Theorem (1)

Assume that there exist a small set  $\mathcal{D}$ , a function  $V : X \to [1, \infty)$ and a continuously differentiable increasing concave function  $\phi : [1, \infty) \to (0, \infty)$ , such that  $\sup_{\mathcal{D}} V < \infty$ ,  $\inf_{[1,\infty)} \phi > 0$ , and

$$PV \leq V - \phi \circ V + b\mathbf{1}_{\mathcal{D}}.$$

Fix  $\beta \in (0,1)$  and let  $n : X \to \mathbb{N}$  satisfy  $n(x) \sim \frac{1}{\beta} \left( \frac{V}{\phi \circ V} \right)(x)$ . Then for any  $\beta < \beta' < 1$ ,

$$P^{n(x)}W \leq \beta'W + b'\mathbf{1}_{\mathcal{C}},$$

where  $W = \phi \circ V$ .

Introduction 000000	Establishing drift O	Using drift 000	Applications	References
Comments				

- Theorem (1) says that we can deduce a (state-dependent) subsampled geometric drift condition from a one-step drift, but on a different scale;
- More general results (not requiring a drift condition for V) can be stated.

Introduction 000000	Establishing drift •	Using drift 000	Applications	References
Examples				

$$PV \leq V - \phi \circ V + b\mathbf{1}_{\mathcal{D}} \qquad \Rightarrow \qquad P^{n(x)}W \leq \beta'W + b'\mathbf{1}_{\mathcal{C}}$$

• Polynomially ergodic: if  $\phi(t) \sim ct^{1-\alpha}$  for some  $\alpha \in (0,1)$ , then  $n \sim V^{\alpha}$  and  $W = V^{1-\alpha}$ .



$$PV \leq V - \phi \circ V + b\mathbf{1}_{\mathcal{D}} \qquad \Rightarrow \qquad P^{n(x)}W \leq \beta'W + b'\mathbf{1}_{\mathcal{C}}$$

• Polynomially ergodic: if  $\phi(t) \sim ct^{1-\alpha}$  for some  $\alpha \in (0,1)$ , then  $n \sim V^{\alpha}$  and  $W = V^{1-\alpha}$ .

*Motivation:* Deterministic chain:  $\Phi_{n+1} = \Phi_n - \Phi_n^{0.4}$  (so V(x) = x)



Introduction 000000	Establishing drift •	Using drift 000	Applications	References
Examples				

$$PV \leq V - \phi \circ V + b\mathbf{1}_{\mathcal{D}} \qquad \Rightarrow \qquad P^{n(x)}W \leq \beta'W + b'\mathbf{1}_{\mathcal{C}}$$

- Polynomially ergodic: if  $\phi(t) \sim ct^{1-\alpha}$  for some  $\alpha \in (0,1)$ , then  $n \sim V^{\alpha}$  and  $W = V^{1-\alpha}$ .
- Subgeometrically ergodic: if  $\phi(t) \sim t[\ln t]^{-\alpha}$  for some  $\alpha > 0$ , then  $n \sim [\ln V(x)]^{\alpha}$  and  $W = V[\ln V]^{-\alpha}$ .

Introduction 000000	Establishing drift •	Using drift 000	Applications	References
Examples				

$$PV \leq V - \phi \circ V + b\mathbf{1}_{\mathcal{D}} \qquad \Rightarrow \qquad P^{n(x)}W \leq \beta'W + b'\mathbf{1}_{\mathcal{C}}$$

- Polynomially ergodic: if  $\phi(t) \sim ct^{1-\alpha}$  for some  $\alpha \in (0,1)$ , then  $n \sim V^{\alpha}$  and  $W = V^{1-\alpha}$ .
- Subgeometrically ergodic: if  $\phi(t) \sim t[\ln t]^{-\alpha}$  for some  $\alpha > 0$ , then  $n \sim [\ln V(x)]^{\alpha}$  and  $W = V[\ln V]^{-\alpha}$ .
- Logarithmically ergodic: if  $\phi(t) \sim [1 + \ln t]^{\alpha}$  for some  $\alpha > 0$ , then  $n \sim \frac{V}{[1 + \ln V]^{\alpha}}$  and  $W = [1 + \ln V]^{\alpha}$ .



Question 2

Suppose we know that

$$P^{n(x)}V(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x).$$

What can be said about moments of  $\tau_{\mathcal{C}}$ ?



Question 2

Suppose we know that

$$P^{n(x)}V(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x).$$

What can be said about moments of  $\tau_{\mathcal{C}}$ ?

• If n(x) = c then  $\Phi$  is geometrically ergodic;



Question 2

Suppose we know that

$$P^{n(x)}V(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x).$$

What can be said about moments of  $\tau_{\mathcal{C}}$ ?

- If n(x) = c then  $\Phi$  is geometrically ergodic;
- Alternative drift condition

$$P^{n(x)}V(x) \leq \beta^{n(x)}V(x) + b\mathbf{1}_{\mathcal{C}}(x)$$

(with no relation assumed between n and V) shown by Meyn & Tweedie (1994) to also imply geometric ergodicity.

Introduction	Establishing drift	Using drift	Applications	References
000000	0	000		

## Theorem (2)

#### Assume that

$$P^{n(x)}V(x) \leq \beta V(x) + b\mathbf{1}_{\mathcal{C}}(x).$$

If there exists a strictly increasing function  $R: (0,\infty) \to (0,\infty)$  satisfying one of the following conditions

- (i)  $t \mapsto R(t)/t$  is non-increasing and  $R \circ n \leq V$ ,
- (ii) R is a convex continuously differentiable function such that R' is log-concave and  $R^{-1}(V) R^{-1}(\beta V) \ge n$ ,

then there exists a constant M such that

$$\mathbb{E}_{x}[R(\tau_{\mathcal{C}})] \leq M(V(x) + b\mathbf{1}_{\mathcal{C}}(x)).$$

Introduction 000000	Establishing drift O	Using drift 000	Applications	References

# • If $n(x) \leq V(x)$ then taking R(t) = t in (i) we obtain

$$\sup_{x\in\mathcal{C}}\mathbb{E}_x[ au_{\mathcal{C}}]<\infty$$
 ;

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References

• If  $n(x) \leq V(x)$  then taking R(t) = t in (i) we obtain

$$\sup_{x\in\mathcal{C}}\mathbb{E}_x[ au_{\mathcal{C}}]<\infty$$
 ;

Part (i) can be applied when V = ξ ∘ n for some increasing concave function ξ (*i.e.* useful when n ≫ V);

Introduction 000000	Establishing drift 0	Using drift 000	Applications	References

• If  $n(x) \leq V(x)$  then taking R(t) = t in (i) we obtain

$$\sup_{x\in\mathcal{C}}\mathbb{E}_x[ au_{\mathcal{C}}]<\infty$$
 ;

- Part (i) can be applied when V = ξ ∘ n for some increasing concave function ξ (*i.e.* useful when n ≫ V);
- Alternatively, if  $n = \xi \circ V$  then

$$R^{-1}(t) \sim \int_1^t rac{\xi(u)}{u} du$$

satisfies (ii) (*i.e.* this is useful when n/V decreasing).



**Geometric rates:** if n(x) = 1, take  $R(t) = \kappa^t$ , with  $1 \le \kappa \le \beta^{-1}$ . Then

• *R* is convex and log-concave;

• 
$$R^{-1}(V) - R^{-1}(\beta V) = (\ln \beta^{-1})/(\ln \kappa) \ge 1 = n.$$

Thus (ii) shows that

 $\mathbb{E}_{x}\left[R(\tau_{\mathcal{C}})\right] = \mathbb{E}_{x}\left[\kappa^{\tau_{\mathcal{C}}}\right] \leq M\left(V(x) + b\mathbf{1}_{\mathcal{C}}(x)\right)\,,$ and so  $\mathbb{E}_{x}\left[\beta^{-\tau_{\mathcal{C}}}\right] < \infty$ .



**Polynomial rates:** suppose  $n(x) \sim V^{\frac{\alpha}{(1-\alpha)}}(x)$ , for some  $\alpha \in (0,1]$ . Letting  $R(t) \sim t^{1/\alpha-1}$ , we see that

- when  $\alpha \leq 1/2$  then R satisfies (ii);
- when  $\alpha \geq 1/2$  then R(t)/t is non-increasing, and

$$R \circ n \sim (V^{\frac{\alpha}{(1-\alpha)}})^{1/\alpha-1} = V$$
,

and so R satisfies (i).



**Polynomial rates:** suppose  $n(x) \sim V^{\frac{\alpha}{(1-\alpha)}}(x)$ , for some  $\alpha \in (0,1]$ . Letting  $R(t) \sim t^{1/\alpha-1}$ , we see that

- when  $\alpha \leq 1/2$  then R satisfies (ii);
- when  $\alpha \geq 1/2$  then R(t)/t is non-increasing, and

$$R \circ n \sim (V^{\frac{\alpha}{(1-\alpha)}})^{1/\alpha-1} = V$$
,

and so R satisfies (i).

In either case, we obtain

$$\mathbb{E}_{x}\left[\tau_{\mathcal{C}}^{1/\alpha-1}\right] \leq cV(x).$$

Logarithmic and subgeometric rates can be dealt with similarly.

Introduction 000000	Establishing drift 0	Using drift 00●	Applications	References
Corollar	y (to Theorem (2))	)		
If $\pi(V)$	$<\infty$ then there ex	kists a small set 1	D with	

$$\sup_{x\in\mathcal{D}}\mathbb{E}_{x}\left[\sum_{k=0}^{\tau_{\mathcal{D}}}R(k)\right]<\infty.$$

Introduction	Establishing drift	Using drift	Applications	References
		000		

## Corollary (to Theorem (2))

If  $\pi(V) < \infty$  then there exists a small set  ${\mathcal D}$  with

$$\sup_{x\in\mathcal{D}}\mathbb{E}_{x}\left[\sum_{k=0}^{\tau_{\mathcal{D}}}R(k)\right]<\infty.$$

**Polynomial chains:** if  $PV \leq V - cV^{1-\alpha} + b\mathbf{1}_{\mathcal{C}}$  then (Thm (1))

$$P^{n(x)}W(x) \leq \beta'W(x) + b'\mathbf{1}_{\mathcal{D}},$$

where  $W \sim V^{1-\alpha}$  and  $n \sim V^{\alpha}$ ; furthermore,  $\pi(W) < \infty$ .

Introduction	Establishing drift	Using drift	Applications	References
		000		

## Corollary (to Theorem (2))

If  $\pi(V) < \infty$  then there exists a small set  ${\mathcal D}$  with

$$\sup_{x\in\mathcal{D}}\mathbb{E}_{x}\left[\sum_{k=0}^{\tau_{\mathcal{D}}}R(k)\right]<\infty.$$

**Polynomial chains:** if  $PV \leq V - cV^{1-\alpha} + b\mathbf{1}_{\mathcal{C}}$  then (Thm (1))

$$P^{n(x)}W(x) \leq \beta'W(x) + b'\mathbf{1}_{\mathcal{D}},$$

where  $W \sim V^{1-\alpha}$  and  $n \sim V^{\alpha}$ ; furthermore,  $\pi(W) < \infty$ . Using  $R(t) \sim t^{1/\alpha - 1}$  in the Corollary we obtain

$$\mathbb{E}_{\mathsf{X}}\left[\tau_{\mathcal{C}}^{1/\alpha}\right] < \infty \,.$$



Class of Markov chains introduced by SBC & Kendall (2007).

#### Definition

 $\Phi$  is *tame* if the following two conditions hold:

(i) there exist  $\delta \in (0, 1)$  and a deterministic function *n* satisfying  $n(x) \le W^{\delta}(x)$  such that

$$\mathbb{E}_{x}\left[W(\Phi_{n(x)})\right] \leq \beta W(x) + b\mathbf{1}_{\mathcal{C}}(x);$$

(ii) the constant  $\delta$  satisfies  $\ln \beta < \delta^{-1} \ln(1-\delta)$ .

*i.e.*  $\Phi$  satisfies a subsampled geometric drift condition, where the subsampling time *n* is not too large.

### Theorem (SBC & Kendall, 2007)

If  $\Phi$  is tame then there exists a perfect simulation algorithm for  $\Phi$  (using Dominated Coupling from the Past).

**Idea of proof:** there exists a simple dominating process for any geometrically ergodic chain (Kendall, 2004); *delay* this (using n) to produce dominating process for X.



### Theorem (SBC & Kendall, 2007)

If  $\Phi$  is tame then there exists a perfect simulation algorithm for  $\Phi$  (using Dominated Coupling from the Past).

**Idea of proof:** there exists a simple dominating process for any geometrically ergodic chain (Kendall, 2004); *delay* this (using n) to produce dominating process for X.



### Theorem (SBC & Kendall, 2007)

If  $\Phi$  is tame then there exists a perfect simulation algorithm for  $\Phi$  (using Dominated Coupling from the Past).

**Idea of proof:** there exists a simple dominating process for any geometrically ergodic chain (Kendall, 2004); *delay* this (using n) to produce dominating process for X.



Introduction 000000	Establishing drift 0	Using drift 000	Applications	References
When is a cl	nain tame?			

• All geometrically ergodic chains are tame.

#### Proposition

If  $PV \leq V - cV^{1-\alpha} + b\mathbf{1}_{\mathcal{C}}$ , with  $\alpha \in (0, 1/2)$ , then  $\Phi$  is tame.

- Proof easy, using Theorem (1).
- Follows that chains with subgeometric drift  $(\phi(t) \sim t[\ln t]^{-\alpha})$  are tame.
- Logarithmically ergodic chains  $(\phi(t) \sim [1 + \ln t]^{\alpha})$  not covered by this result.

Introduction 000000	Establishing drift O	Using drift 000	Applications	References
Dominating	process			

Can also use above theory to determine ergodic properties of the dominating process D for  $\Phi$  in the perfect simulation algorithm.

- D does not satisfy a simple one-step drift condition;
- but establishing state-dependent drift is simple!
- Theorem (2) provides information about ergodic properties of *D*.



Introduction 000000	Establishing drift O	Using drift 000	Applications	References
Dominating	process			

Can also use above theory to determine ergodic properties of the dominating process D for  $\Phi$  in the perfect simulation algorithm.

- D does not satisfy a simple one-step drift condition;
- but establishing state-dependent drift is simple!
- Theorem (2) provides information about ergodic properties of *D*.

*E.g.* if  $n(x) \sim W(x)^{\gamma}$  (with  $\gamma \leq \delta$ ), then *D* is ergodic and converges to  $\pi_D$  polynomially fast (in total variation)

• but note that this isn't enough to guarantee that the mean run-time of the domCFTP algorithm is finite ...



- Sufficient conditions for ergodicity of strong Markov processes
  - applications in queueing and network stability
  - continuum range of rates of convergence
  - explicit norm of convergence





- Sufficient conditions for ergodicity of strong Markov processes
  - applications in queueing and network stability
  - continuum range of rates of convergence
  - explicit norm of convergence



• Yüksel & Meyn (2012) use *random-time, state-dependent drift criteria* to prove stability results, but *not* convergence rates . . .

Introduction 000000	Establishing drift O	Using drift 000	Applications	References
References				

#### Connor, S. B. and G. Fort (2009).

State-dependent FosterLyapunov criteria for subgeometric convergence of Markov chains.

Stochastic Processes and their Applications 119(12), 4176–4193.

Connor, S. B. and W. S. Kendall (2007). Perfect Simulation for a Class of Positive Recurrent Markov Chains. *Ann. Appl. Probab.* 17, 781–808.

Kendall, W. S. (2004).

Geometric Ergodicity and Perfect Simulation.

Meyn, S. P. and R. L. Tweedie (1994).

State-dependent criteria for convergence of Markov chains.

Ann. Appl. Probab. 4, 149-168.

#### Yüksel, S. and S. Meyn (2012).

Random-Time, State-Dependent Stochastic Drift for Markov Chains and Application to Stochastic Stabilization Over Erasure Channels. *Preprint*.