ODE methods for Markov chain stability with applications to MCMC *

G. Fort[1] E. Moulines[1] S. Meyn[2] P. Priouret[3]

[1]Signal and Image Processing Lab., LTCI UMR5141, Ecole Nationale Supérieure des Télécommunications

[2]Coordinated Sience Lab., University of Illinois

[3]Laboratoire de Probabilité, Université Paris VI

New Developments in MCMC, Warwick, 2006

---

*as announced in the program!

Outline of the talk

- Motivations
- Fluid Limits: Construction and Main Results
- Applications to MCMC
- Conclusion & Discussion

The Metropolis Algorithm

- The Metropolis algorithm is a popular method for generating samples from virtually any distribution π.
- Idea: simulate a Markov chain $\{\Phi_k\}_{k\geq 0}$ on X with transition kernel P where P is such that $\pi P = \pi$.
- Algorithm:
  1. Given that the chain is at x, draw y from some symmetric distribution q.
  2. With probability $\alpha(x, x+y) = 1 \wedge \frac{\pi(x+y)}{\pi(x)}$, the chain moves to $x+y$. Otherwise, the chain stays at its current location x.
- P is reversible with respect to π, and therefore admits π as invariant distribution.

Stability and Rate of Convergence

- Questions:
  - Stability ? for which functions can I expect that $\mathbb{E}_x[f(\Phi_k)] \to \pi(f)$ for any $x \in \mathsf{X}$?
  - Rate of Convergence ? Can I determine a sequence $\{r(n)\}$ and control functions f and V such that, for any $x \in \mathsf{X}$,

$$r(n) \sup_{|\phi| \leq f} |P^n f(x) - \pi(f)| \leq V(x)$$

Problems

- Classical answers: identify a small set C and a drift function V satisfying either
  - Foster-Lyapunov drift conditions, e.g. $PV \leq \lambda V + b\mathbb{1}_C$
  - (Gersende Fort) Weaker form drift, e.g. $PV \leq V - cV^\alpha + b\mathbb{1}_C$.
- Problem: Identifying small sets for MCMC is not a difficult issue (but minorizing constants can be very small !)... but checking drift conditions is often a tedious job.

- Today: another approach inspired by fluid limit techniques developed to study the stability of stochastic networks

The Metropolis Algorithm: Martingale Decomposition

$$\Phi_{k+1} = \Phi_k + B_{k+1} Y_{k+1} \qquad B_k \sim \mathsf{Ber}\left(\alpha(\Phi_k, \Phi_k + Y_{k+1})\right)$$
$$= \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$$

1. $\epsilon_{k+1} \stackrel{\text{def}}{=} \Phi_{k+1} - \mathbb{E}\left[\Phi_{k+1} \mid \Phi_k\right]$ is a martingale increment
2. $\Delta(x) = \int_{\mathsf{R}_x} y \left(\frac{\pi(x+y)}{\pi(x)} - 1\right) q(y) dy$ where

   $\mathsf{R}_x \stackrel{\text{def}}{=} \{y, \pi(x+y) < \pi(x)\}$ is the (potential) rejection region.

Can we say something sensible from this simple recurrence equation ?

The classical ODE Method

- Consider a dynamical system described by the recurrence equation

$$\Phi^{\gamma}_{k+1} = \Phi^{\gamma}_{k} + \gamma H(\Phi^{\gamma}_{k}, U_{k+1}) \quad , k \geq 0 \, ,$$

where $\{U_k\}$ is an i.i.d. † sequence, $H : \mathsf{X} \times \mathbb{R} \to \mathsf{X}$ is a smooth function and γ is a small parameter.

- ODE method: guess properties of $\{\Phi^{\gamma}_{k}\}$ from the ODE

$$\dot{\mu} \stackrel{\text{def}}{=} h(\mu)$$

where $h(x) \stackrel{\text{def}}{=} \mathbb{E}[H(x, U)]$ is the mean field.

---

† More complicated noise models can be considered

ODE method: characterization of the stationary distributions

(after Fredlin-Wentzell, Fort-Pages)

- $\{\Phi_k^\gamma\}$ is an homogeneous Markov chain.
- Under mild conditions, for $0 < \gamma \leq \gamma_0$, the chain has (at least) one invariant distribution π^γ. In addition, the set $\{\pi^\gamma\}_{\gamma \leq \gamma_0}$ is tight.
- The ODE $\dot{\mu} = h(\mu)$ has invariant distribution(s) satisfying the flow-invariance property. Denote by $\mathcal{J}_h$ the set of such distributions.
- Results The limiting points of $\{\pi^{\gamma_n}\}$ where $\{\gamma_n\}$ is any sequence satisfying $\lim_n \gamma_n = 0$ is included in $\mathcal{J}_h$.

ODE method and Stochastic Stability

(after Borkar-Meyn)

- ODE can also be used to establish stochastic stability of

$$\begin{aligned}\Phi_{k+1}^{\gamma} &= \Phi_k^{\gamma} + \gamma H(\Phi_k^{\gamma}, U_{k+1}) \\ &= \Phi_k^{\gamma} + \gamma h(\Phi_k^{\gamma}) + \gamma \epsilon_{k+1}^{\gamma}\end{aligned}$$

where $\epsilon_{k+1}^{\gamma} \stackrel{\text{def}}{=} H(\Phi_k^{\gamma}, U_{k+1}) - h(\Phi_k^{\gamma})$.

- Assumptions:
  1. the radial limits $h_\infty(\mu) = \lim_{r\to\infty} h(r\mu)/r$ exist and the origin is an asymptotically stable equilibrium for the limiting ODE $\dot{\mu} = h_\infty(\mu)$
  2. $\mathbb{E}\left[|\epsilon_{k+1}^{\gamma}|^p \,\middle|\, \mathcal{F}_k \right] \le C(1 + |\Phi_k^{\gamma}|^p)$.
- Results: The Markov chain $\{\Phi_k\}$ is stable in the sense that, for all $0 \le \gamma \le \gamma_0$, i.e. for any $x \in \mathsf{X}$,

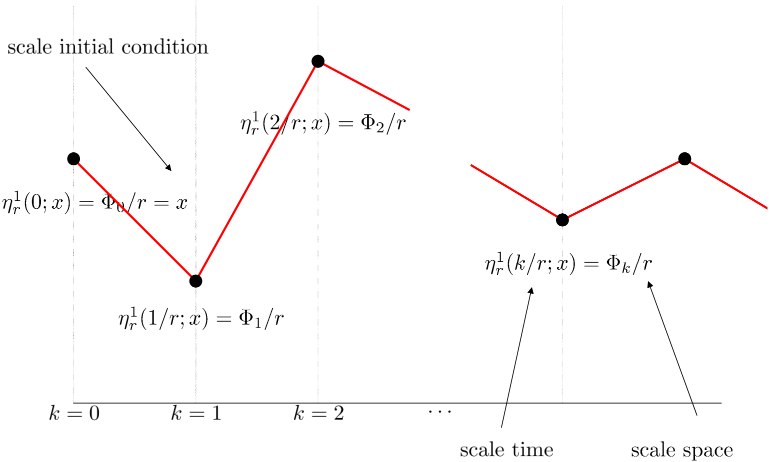
$$\limsup_{n\to\infty} \mathbb{E}_x\left(|\Phi_k^{\gamma}|^p \right) < \infty$$

The ODE method can be used to establish both the stability and convergence of the stochastic approximation method.

Interpolated Process

- In the ODE method, all the results are obtained by letting the stepsize $\gamma \to 0$.
- In the recursion $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$ there is no such small parameter...
- To be able to play with limits (Gareth's talk), scale $\{\Phi_k\}$ simultaneously in SPACE, TIME and INITIAL CONDITION.

Interpolated Process



Understanding the scaling

Rewrite the recurrence equation as a function of the interpolated process

$$\eta_r^1\left((k+1)/r;x\right) = \eta_r^1\left(k/r;x\right) + r^{-1}\Delta\left(r\eta_r^1\left(k/r;x\right)\right) + r^{-1}\epsilon_{k+1}\ .$$

Assuming the existence of the radial limits $\lim_{r\to\infty}\Delta(r\mu) = h(\mu)$

$$\eta_r^1\left(t;x\right) = \eta_r^1\left(0;x\right) + \int_0^t h\left(\eta_r^1\left(u;x\right)\right)du + E_r^1(t;x) + R_r^1(t;x)\ ,$$

where $E_r^1(t;x) \stackrel{\text{def}}{=} r^{-1}\sum_{k=0}^{[tr]}\epsilon_{k+1}$ and $R_r^1(t;x)$ is a remainder term.

Understanding the scaling ($\alpha = 1$)

- Provided that $\sup_x \mathbb{E}_x[|\epsilon_1|^p] < \infty$, then

$$\sup_x \mathbb{E}_x \left| r^{-1} \sum_{k=0}^{[tr]} \epsilon_{k+1} \right|^p \leq C r^{-p} (tr)^{p/2}$$

  which goes to zero.
- Therefore, if we can control the error $|\Delta(r\mu) - h(\mu)|$ as $r \to \infty$, it is not difficult to believe that the sequence $\{\eta_r^1(t;x)\}$ converges to a limit $\{\eta_\infty^1(t;x)\}$ which happens to be the solution of the ODE $\dot{\mu} = h(\mu)$...

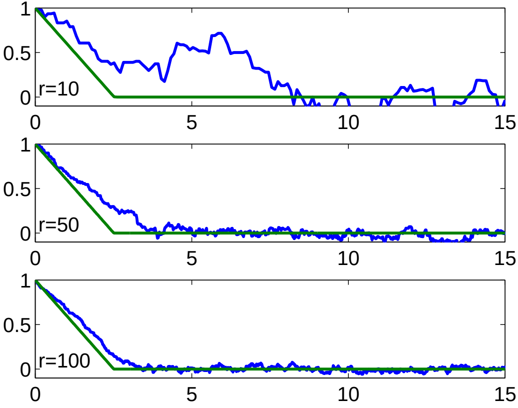
Example: Metropolis Algorithm on $\mathbb{R}$

π	$h(x) = \lim_{r\to\infty} \Delta(rx)$
$e^{-x^2/2\sigma^2}$	$-\text{sign}(x)\ C_G$
$e^{-\|x\|/\sigma}$	$-\text{sign}(x)\ C_L$
$e^{-\|x/\sigma\|^\beta}$ $0 < \beta < 1$	0

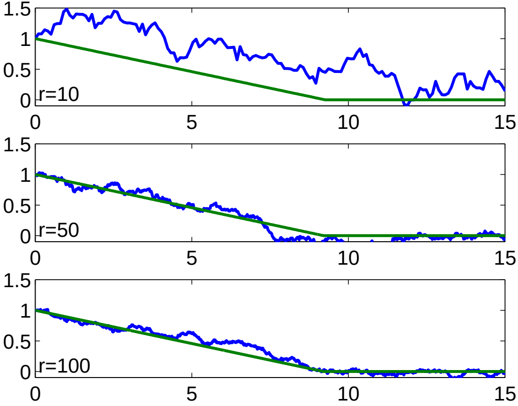
Table: ODE associated for the Metropolis Algorithm for different choices of the target distribution

Note that the field is radially invariant, $h(rx) = h(x)$

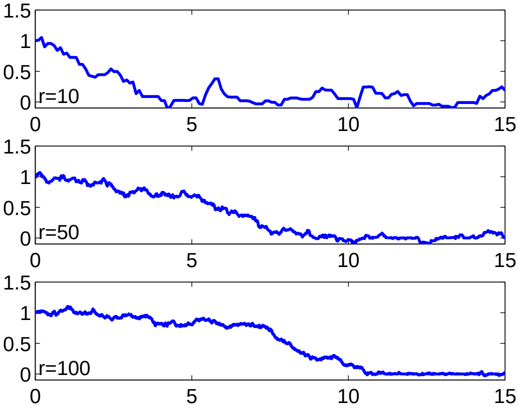
Metropolis on $\mathbb{R}$: Gaussian Target



Metropolis on $\mathbb{R}$: Laplacian Target



Metropolis on $\mathbb{R}$: Weibulian Target



Questions

1. Can we relax the assumptions ?
   - What happens if $\Delta(rx)$ do not have radial limits for all x ?
   - How can we avoid that the mean field $h \equiv 0$ be trivial ?
2. What does the stability of the solutions of the ODE (or more generally, of potential limits of the scaled process) tell us about the stability of the Markov chain $\{\Phi_k\}$?

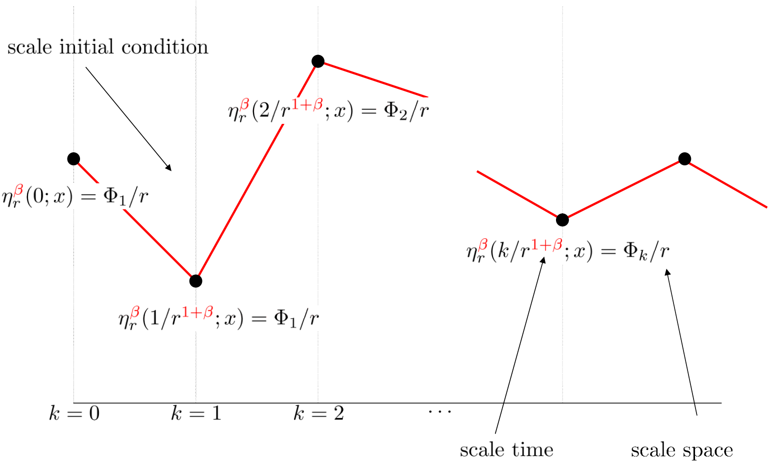
Changing the normalization

- For the Weibul target, we do not obtain a meaningful limit because $\lim_{r\to\infty} \Delta(rx) = 0$...
- Nevertheless, modifying the normalization,

$$\lim_{r\to\infty} |x|^{1-\delta} r^{1-\delta} \Delta(rx) = -\mathrm{sign}(x) C_W \stackrel{\mathrm{def}}{=} \Delta_\infty(x) \ ,$$

  where $C_W \stackrel{\mathrm{def}}{=} \delta \int_0^\infty y^2 q(y) dy$.
- Solution: Use a different scaling for TIME and SPACE.
- Set $\beta = 1 - \delta$ and consider the polygonal process $\eta_r^\beta(t; x)$ that agrees with Φ_k / r at the knots $k/r^{1+\beta}$.

Changing the normalization



Changing the normalization

$$\eta_r^\beta(t;x) = \eta_r^\beta(0;x) + \int_0^t h(\eta_r^\beta(u;x))du + E_r^\beta(t;x) + R_r^\beta(t;x)$$

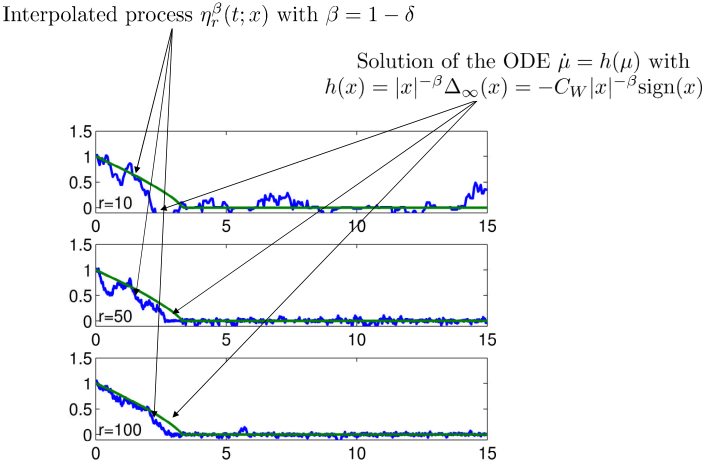
where $h(x) \stackrel{\text{def}}{=} |x|^{-\beta}\Delta(x)$, $E_r^\beta(t;x) \stackrel{\text{def}}{=} r^{-1}\sum_{k=1}^{\lceil tr^{1+\beta}\rceil} \epsilon_k$ and $R_r^\beta(t;x)$ is a remainder term.

- Provided that $\sup_x \mathbb{E}_x[|\epsilon_1|^p] < \infty$, then

$$\sup_x \mathbb{E}_x|E_r^\beta(t;x)|^p \leq Cr^{-p}(tr)^{p(1+\beta)/2}$$

  which goes to zero as $r \to \infty$ provided that $\beta < 1$.
- When $\beta = 1$, the limit are no longer deterministic (work in progress)

Metropolis on $\mathbb{R}$: weibulian target



Tightness and Continuity

- Denote by $\mathsf{C}(\mathsf{I}, \mathsf{X})$ the space of continuous X-valued functions on $\mathsf{I} \subseteq \mathbb{R}^+$ equipped with the uniform topology.
- Denote by $\mathbb{Q}^\alpha_{r;x}$ the distribution of $\eta^\alpha_r(\cdot; x)$ on $\mathsf{C}(\mathbb{R}^+, \mathsf{X})$.
- A probability measure $\mathbb{Q}^\alpha_x$ on $\mathsf{C}(\mathbb{R}^+, \mathsf{X})$ is said to be an α-fluid limit if there exist $\{r_n\} \subset \mathbb{R}_+$ and $\{x_n\} \subset \mathsf{X}$ satisfying $\lim_{n\to\infty} r_n = +\infty$ and $\lim_{n\to\infty} x_n = x$ such that $\{\mathbb{Q}^\alpha_{r_n;x_n}\}$ converges weakly to $\mathbb{Q}^\alpha_x$ on $\mathsf{C}(\mathbb{R}^+, \mathsf{X})$

$$\mathbb{Q}^\alpha_{r_n;x_n} \Rightarrow \mathbb{Q}^\alpha_x .$$

Tightness and Continuity

- Assumption: $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$ with $\{\epsilon_k\}$ martingale increment and
  1. $\lim_{K\to\infty} \sup_{x\in\mathsf{X}} \mathbb{E}_x[|\epsilon_1|^p \mathbb{1}\{|\epsilon_1| \geq K\}] = 0$
  2. $\sup_{x\in\mathsf{X}} \left\{(1+|x|^\beta)|\Delta(x)|\right\} < \infty$
- Result: For all $0 \leq \alpha \leq \beta$ and any sequences satisfying $\lim_{n\to\infty} r_n = +\infty$ and $\lim_{n\to\infty} x_n = x$,

$$\{\mathbb{Q}^\alpha_{r_n;x_n}\} \text{ is tight}$$

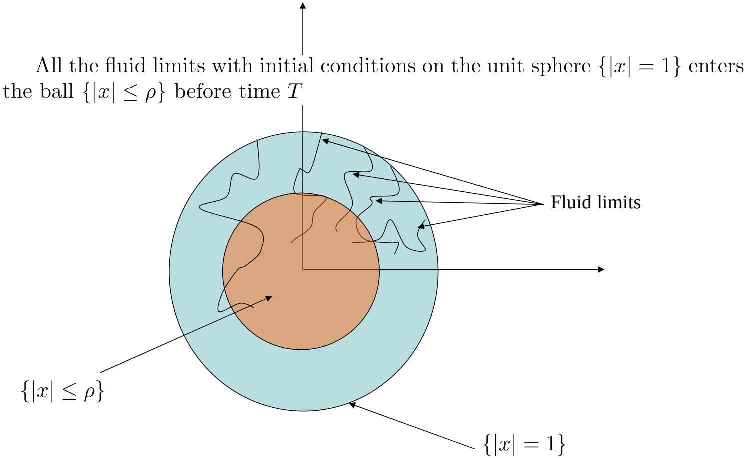
- Fluid limits exist even if Δ does not have meaningful radial limits!
- Any limiting point $\mathbb{Q}^\alpha_x$ of a sequence $\{\mathbb{Q}^\alpha_{r_n;x_n}\}$ is referred to as an α-fluid limit and the set of all possible α-fluid limits is called the α-fluid limit model.

Stability of the fluid limit model

Stability (after Stolyar): The β-fluid limit model is said to be stable if there exist $T > 0$ and $\rho < 1$ such that for any $x \in \mathsf{X}$ with $|x| = 1$,

$$\mathbb{Q}_x^\beta \left(\eta \in \mathsf{D}(\mathbb{R}_+, \mathsf{X}), \inf_{0 \le t \le T} |\eta(t)| \le \rho \right) = 1 \; .$$

- Gaussian & Laplacian cases: satisfied with $\beta = 1$.
- Weibulian case: satisfied with $\beta = 1 - \delta$.


All the fluid limits with initial conditions on the unit sphere $\{|x| = 1\}$ enters the ball $\{|x| \leq \rho\}$ before time T
Fluid limits
$\{|x| \leq \rho\}$
$\{|x| = 1\}$

Stability of the fluid limit model implies the ergodicity of the Markov Chain

- **Assumptions**:
  1. $\{\Phi_k\}$ phi-irreducible, aperiodic, and compact sets are petite.
  2. $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$ with $\{\epsilon_k\}$ martingale increment and
     - $\lim_{K\to\infty} \sup_{x\in\mathsf{X}} \mathbb{E}_x[|\epsilon_1|^p \mathbb{1}\{|\epsilon_1| \geq K\}] = 0$
     - $\sup_{x\in\mathsf{X}} \left\{(1+|x|^\beta)|\Delta(x)|\right\} < \infty$
  3. The β-fluid limit model is stable.
- **Results**: for any $1 \leq q \leq (1+\beta)^{-1}p$ and any function f such that $\sup_{x\in\mathsf{X}} |f(x)|/(1+|x|^{p-q(1+\beta)}) < \infty$,

$$n^{q-1} \left|\mathbb{E}_x[f(\Phi_n)] - \pi(f)\right| \to 0$$

Step 1: State Dependent Drift Conditions

- **Assumptions**:
  1. $\{\Phi_k\}_{k\geq 0}$ is a phi-irreducible and aperiodic.
  2. There exist a function $V : \mathsf{X} \to [1, \infty)$, a stopping time $\tau \geq 1$, a constant $\varepsilon \in (0, 1)$ and a petite set $C \subset \mathcal{X}$, such that,

$$\mathbb{E}_x\left[\sum_{k=0}^{\tau-1} V^{1-\alpha}(\Phi_k)\right] \leq V(x)\ , \qquad x \notin C$$

$$\mathbb{E}_x\left[V(\Phi_\tau)\right] \leq (1-\varepsilon)V(x)\ , \qquad x \notin C$$

$$\sup_C\{V + PV\} < \infty\ .$$

- **Results**: Then P is positive Harris with invariant probability π and
  1. $\lim_{n\to\infty} n^{(1-\alpha)/\alpha} \|P^n(x,\cdot) - \pi\|_{\mathrm{TV}} = 0$.
  2. $\lim_{n\to\infty} \|P^n(x,\cdot) - \pi\|_{V^{1-\alpha}} = 0$.

Step 2: From Fluid Limit Model Stability to State-Dependent Drift

- The stability of the fluid limit model implies (by the Portmanteau Theorem)

$$\limsup_{|x|\to\infty} \mathbb{P}_x\left(\sigma > \lceil T|\Phi_0|^{1+\beta}\rceil\right) = 0 ,$$

  with $\sigma \stackrel{\text{def}}{=} \inf\{k \geq 0, |\Phi_k| < \rho|\Phi_0|\}$.
- On the other hand, from $\sup(1+|x|^\beta)|\Delta(x)| < \infty$ and L^p-uniform integrability of ϵ,

$$\sup_{x\in\mathsf{X}}(1+|x|)^{-p}\mathbb{E}_x\left[\sup_{0\leq k\leq\lfloor T|\Phi_0|^{1+\beta}\rfloor}|\Phi_k|^p\right] < \infty ,$$

- Finally, set $\tau \stackrel{\text{def}}{=} \sigma \wedge \lceil T|\Phi_0|^{1+\beta}\rceil$. Combining the two previous results,
  1. there exists M such that $\sup_{|x|\geq M}|x|^{-p}\mathbb{E}_x\left[|\Phi_\tau|^p\right] < 1$,
  2. $\mathbb{E}_x\left[\sum_{k=0}^{\tau-1}|\Phi_k|^p\right] \leq C\,|x|^{p+1+\beta}$

Fluid Limit Characterization

- How can we establish the stability of the β-fluid model ?
- Answer: Not straightforward in general, except when Δ satisfies scaling properties:

Radial Limits

There exist an open cone $\mathsf{O} \subseteq \mathsf{X} \setminus \{0\}$ and a continuous function $\Delta_\infty : \mathsf{O} \to \mathsf{X}$ such that, for any compact subset $\mathsf{H} \subseteq \mathsf{O}$,

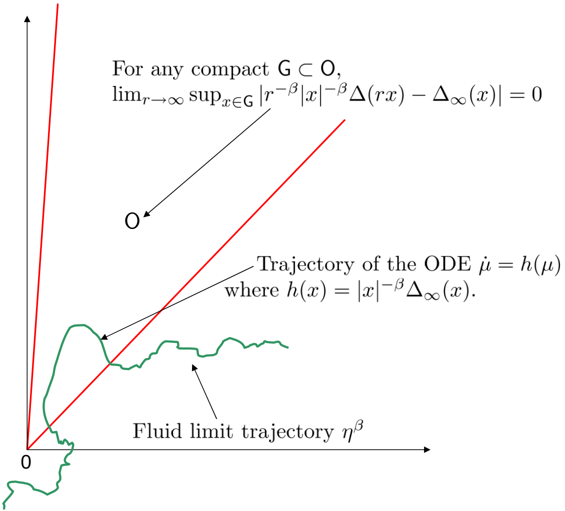
$$\lim_{r \to +\infty} \sup_{x \in \mathsf{H}} \left| r^\beta |x|^\beta \Delta(rx) - \Delta_\infty(x) \right| = 0 \, ,$$

Fluid Limits and ODE flow

- Assumptions:
  1. $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$ with $\{\epsilon_k\}$ martingale increment and
     - $\lim_{K\to\infty} \sup_{x\in\mathsf{X}} \mathbb{E}_x[|\epsilon_1|^p \mathbb{1}\{|\epsilon_1| \geq K\}] = 0$
     - $\sup_{x\in\mathsf{X}} \left\{(1+|x|^\beta)|\Delta(x)|\right\} < \infty$
  2. For any compact set $\mathsf{H} \subset \mathsf{O}$, $\lim_{r\to+\infty} \sup_{x\in\mathsf{H}} \left| r^\beta |x|^\beta \Delta(rx) - \Delta_\infty(x)\right| = 0$,
- Results For any $0 \leq s \leq t$, and any β-fluid limit $\mathbb{Q}_x^\beta$, on the event $\{\eta \in \mathsf{C}(\mathbb{R}^+, \mathsf{X}) : \eta(u) \in \mathsf{O} \text{ for all } u \in [s,t]\}$, the fluid limit agrees with the flow of the ODE $\dot{\mu} = h(\mu)$, where $h(x) \stackrel{\text{def}}{=} |x|^{-\beta}\,\Delta_\infty(x)$, i.e.

$$\sup_{s\leq u\leq t} \left| \eta(u) - \eta(s) - \int_s^u h\circ\eta(v)dv \right| = 0\,, \quad \mathbb{Q}_x^\beta - \text{a.s.}$$

Fluid Limits and ODE flow

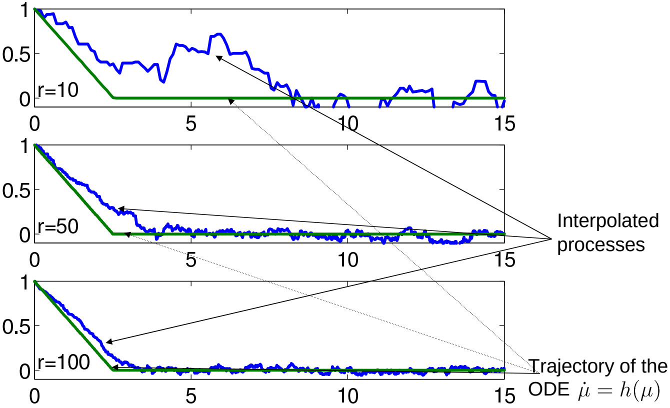


Characterization of the fluid limits

- Assumptions:
  1. $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$ with $\{\epsilon_k\}$ martingale increment and
     - $\lim_{K\to\infty} \sup_{x\in\mathsf{X}} \mathbb{E}_x[|\epsilon_1|^p \mathbb{1}\{|\epsilon_1| \geq K\}] = 0$
     - $\sup_{x\in\mathsf{X}} \left\{(1+|x|^\beta)|\Delta(x)|\right\} < \infty$
  2. for any $x \in \mathsf{X} \setminus \{0\}$ the ODE $\dot{\mu} = h(\mu)$ with initial condition x has a unique solution, denoted $\mu(\cdot; x)$ on an interval $[0, T_x]$.
- Results
  - all β-fluid limits are deterministic and solve the ODE $\dot{\mu} = h(\mu)$.
  - For any $\epsilon > 0$ and $x \in \mathsf{X}$, and any sequences $\{r_n\} \subset \mathbb{R}_+$ and $\{x_n\} \subset \mathsf{X}$ such that $\lim_{n\to\infty} r_n = +\infty$ and $\lim_{n\to\infty} x_n = x$,

$$\lim_n \mathbb{P}_{r_n x_n} \left(\sup_{0\leq t\leq T_x} \left|\eta^\beta_{r_n}(t; x_n) - \mu(t; x)\right| \geq \epsilon \right) = 0 \ .$$

Characterization of the fluid limits



Back to stability

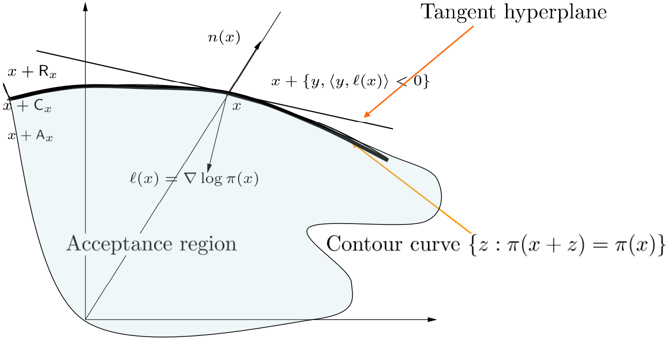
- Assumptions:
  1. $\{\Phi_k\}_{k\in\mathbb{N}}$ phi-irreducible, aperiodic, compact sets are petite.
  2. $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$ with $\{\epsilon_k\}$ martingale increment and
     - $\lim_{K\to\infty} \sup_{x\in\mathsf{X}} \mathbb{E}_x[|\epsilon_1|^p \mathbb{1}\{|\epsilon_1| \geq K\}] = 0$
     - $\sup_{x\in\mathsf{X}} \left\{(1+|x|^\beta)|\Delta(x)|\right\} < \infty$
  3. For all $\mathsf{H} \subset \mathsf{X} \setminus \{0\}$, $\lim_{r\to+\infty} \sup_{x\in\mathsf{H}} \left|r^\beta |x|^\beta \Delta(rx) - \Delta_\infty(x)\right| = 0$,
  4. stability of the ODE solution For some $\rho < 1$ and $T > 0$, $\inf_{[0,T\wedge T_x]} |\mu(\cdot;x)| \leq \rho$ for any $|x| = 1$.
- Results: the β-fluid limit model is stable

Super-Exponential Target

- A probability density function π is said to be super-exponential if π is positive, has continuous first derivatives, and $\lim_{|x|\to\infty}\langle n(x), \ell(x)\rangle = -\infty$ where $\ell(x) \stackrel{\text{def}}{=} \nabla \log \pi(x)$.
- The condition implies that for any $H > 0$ there exists $R > 0$ such that

$$\frac{\pi(x + an(x))}{\pi(x)} \leq \exp(-aH) \quad \text{for } |x| \geq R, a \geq 0 \ ,$$

that is, $\pi(x)$ is at least exponentially decaying along any ray with the rate H tending to infinity as $|x|$ goes to infinity.



Tangent hyperplane
$n(x)$
$x + \mathsf{R}_x$
$x + \{y, \langle y, \ell(x)\rangle < 0\}$
$x + \mathsf{C}_x$
x
$x + \mathsf{A}_x$
$\ell(x) = \nabla \log \pi(x)$
Acceptance region
Contour curve $\{z : \pi(x+z) = \pi(x)\}$

Radial Limits and Homogeneity

- Assumption: The family of rejection regions $\{\mathsf{R}_{rx}, r \geq 0, x \in \mathsf{O}\}$ has radial limits over $\mathsf{O} \subseteq \mathsf{X} \setminus \{0\}$ if there exists $\{\mathsf{R}_{\infty,x}, x \in \mathsf{O}\}$ s.t., for any compact subset $\mathsf{H} \subseteq \mathsf{O}$,

$$\lim_{r \to \infty} \sup_{x \in \mathsf{H}} Q\left(\mathsf{R}_{rx} \ominus \mathsf{R}_{\infty,x}\right) = 0$$

  .

- Result: for any compact set $\mathsf{H} \subset \mathsf{O}$, $\lim_{r \to \infty} \sup_{x \in \mathsf{H}} |\Delta(rx) - \Delta_\infty(x)| = 0$, where

$$\Delta_\infty(x) \stackrel{\text{def}}{=} -\int_{\mathsf{R}_{\infty,x}} y q(y) dy \ .$$

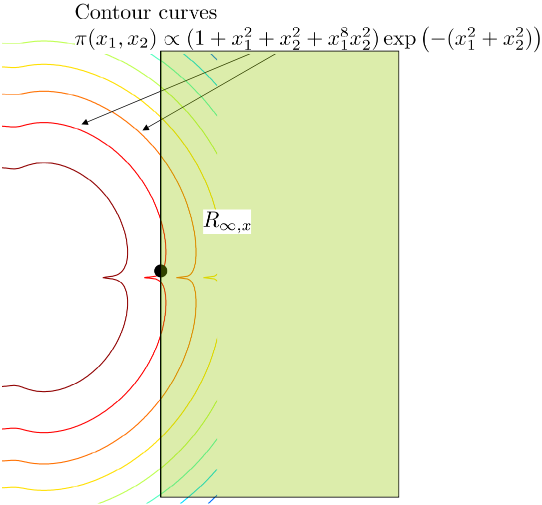
Regularity in the tails

- π is regular in the tails over O if $\{\mathsf{R}_{rx}, r \geq 0, x \in \mathsf{O}\}$ has radial limits over $\mathsf{O} \subseteq \mathsf{X} \setminus \{0\}$ and there exists a continuous function $\ell_\infty : \mathsf{X} \setminus \{0\} \to \mathsf{X}$ such that, for all $x \in \mathsf{O}$,

$$Q\left(\mathsf{R}_{\infty,x} \ominus \{y \in \mathsf{X}, \langle y, \ell_\infty(x)\rangle < 0\}\right) = 0 \; .$$

- Regularity in the tails holds with $\ell_\infty(x) = \lim_{r\to\infty} n(\ell(rx))$ when the curvature at 0 of the contour manifold C_{rx} goes to zero as $r \to \infty$.
- nevertheless, this regularity in the tails still holds in some situations where the curvature of the contour manifolds grow to infinity (along manifolds).

Regularity in the tails



Fluid Limit for super-exponential densities

- Assumption: π is *super-exponential* and *regular in the tails* over $\mathsf{O} \subseteq \mathsf{X} \setminus \{0\}$. The proposal q is rotationally invariant and has bounded moment of order $p > 1$.
- Results:
  1. phi-irreducible, aperiodic, compact sets are small,
  2. $\Phi_{k+1} = \Phi_k + \Delta(\Phi_k) + \epsilon_{k+1}$, with $\sup_{x \in \mathsf{X}} |\Delta(x)| < \infty$ and $\{\epsilon_k\}$ martingale increment satisfying $\sup_{x \in \mathsf{X}} \mathbb{E}_x |\epsilon_1|^p < \infty$.
  3. For any compact subset $\mathsf{H} \subset \mathsf{O}$, $\lim_{r \to \infty} \sup_{x \in \mathsf{H}} |\Delta(rx) - \Delta_\infty(x)| = 0$ with

$$\Delta_\infty(x) = m_1(q_0) \frac{\ell_\infty(x)}{|\ell_\infty(x)|} \ ,$$

where $m_1(q_0) \stackrel{\text{def}}{=} \int_{\mathsf{X}} y_1 \mathbb{1}_{\{y_1 \geq 0\}} q_0(y) dy > 0$, where $y = (y_1, \ldots, y_d)$.

The ODE may be seen as a version of steepest ascent algorithm to maximize $\log \pi$!

Fluid limit for the exponential family

- The tail regularity condition and the definition of the ODE limit are more transparent in a class of models which are very natural in many statistical contexts, namely, the exponential family.
- Define the class $\mathcal{P}$ to consist of those everywhere positive densities with continuous second derivatives π satisfying

$$\pi(x) \propto g(x) \exp\{-p(x)\}$$

  - g is a positive function slowly varying at infinity,
  - p is a positive polynomial in X of even order m and $\lim_{|x|\to\infty} p_m(x) = +\infty$, where p_m denotes the polynomial consisting only of the p's m-th order terms.

Fluid limit for the exponential family

- Assumptions $\pi \in \mathcal{P}$ and let q is rotationally invariant.
- Results
  - π is super-exponential,
  - π is regular in the tails over $\mathsf{X} \setminus \{0\}$ with $\ell_\infty(x) = -n\left[\nabla p_m\left(n(x)\right)\right]$.
  - For any $x \in \mathsf{X} \setminus \{0\}$, there exists $T_x > 0$ such that the ODE $\dot{\mu} = m_1(q_0)n(\ell_\infty(\mu))$ with initial condition x has a unique solution on $[0, T_x)$ and $\lim_{t \to T_x^-} \mu(t; x) = 0$.
  - In addition, the fluid limit $\mathbb{Q}_x^0$ is deterministic on $\mathsf{D}([0, T_x], \mathsf{X})$, with support function $\mu(\cdot; x)$.

Regularity in the tails

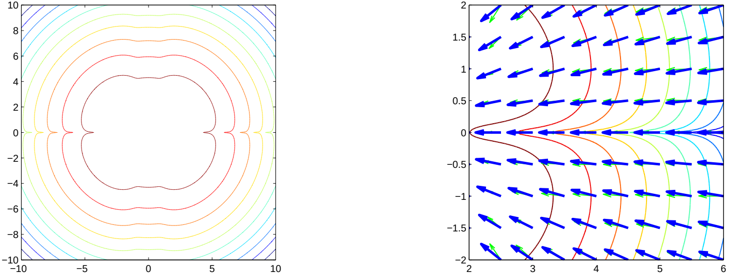
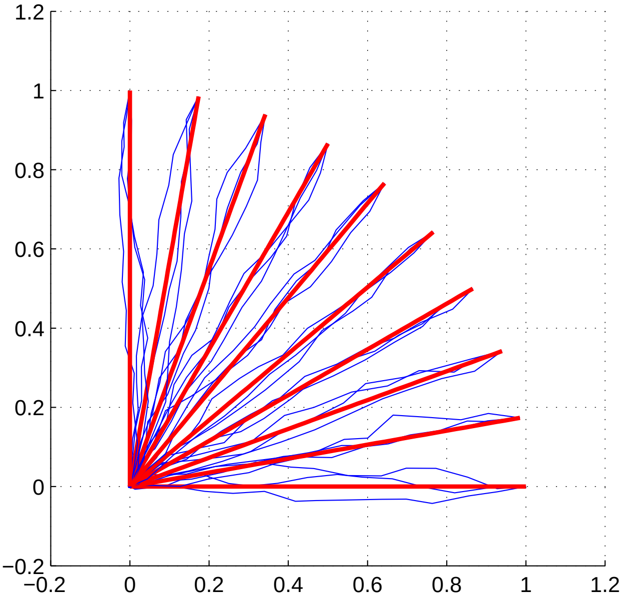


Figure: left panel: $\pi(x_1, x_2) \propto (1 + x_1^2 + x_2^2 + x_1^8 x_2^2) \exp\left(-(x_1^2 + x_2^2)\right)$; right panel: Δ (green arrows) and Δ_∞ (blue arrows)


1.2
1
0.8
0.6
0.4
0.2
0
−0.2
−0.2
0
0.2
0.4
0.6
0.8
1
1.2

The singular case

- In the regular case, we have assumed that Δ has radial limits in every direction of the space, *i.e.* for any compact set $\mathsf{H} \subset \mathsf{O} = \mathsf{X} \setminus \{0\}$, $\lim_{r\to\infty} \sup_{x\in\mathsf{H}} |\Delta(rx) - \Delta_\infty(x)| = 0$.
- What happens if this condition is satisfied only on $\mathsf{O} \subsetneq \mathsf{X} \setminus \{0\}$? What can of conclusions can be reached in this case ?
- The problem becomes then much more involved, and we have got only partial answers (and most presumably, there are not much that can be said in full generality !)
- A special case: the singular regions (the points where there is no radial limit !) are repulsive

The singular case

Theorem

- $\{\Phi_k\}_{k\in\mathbb{N}}$ *phi-irreducible, aperiodic, compact sets are petite, skip-free*
- *radial limits: for all* $\mathsf{H} \subset \mathsf{O} \subsetneq \mathsf{X}$,
  $\lim_{r\to+\infty} \sup_{x\in\mathsf{H}} \left| r^\beta |x|^\beta \Delta(rx) - \Delta_\infty(x) \right| = 0,$

Assume in addition that

- *there exists* $T_0 > 0$ *such that, for any* x, $|x| = 1$, *and any* β*-fluid limit* $\mathbb{Q}_x^\beta$, $\mathbb{Q}_x^\beta\left(\eta : \eta([0, T_0]) \cap \mathsf{O} \neq \emptyset\right) = 1$.
- *for any* $K > 0$, *there exist* $T_K > 0$ *and* $0 < \rho_K < 1$ *such that for any* $x \in \mathsf{O}$, $|x| \leq K$, $\inf_{[0,T_K\wedge T_x]} |\mu(\cdot;x)| \leq \rho_K$
- *for any compact set* $\mathsf{H} \subset \mathsf{O}$ *and any* K,
  $\Omega_\mathsf{H} \stackrel{\text{def}}{=} \{\mu([0, T_x \wedge T_K]; x), x \in \mathsf{H}\}$ *is a compact subset of* O.

Then, the β*-fluid model is stable.*

Mixture of Gaussian distributions

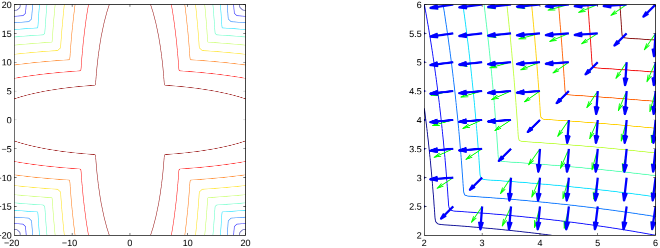
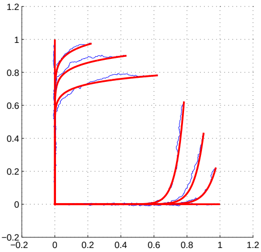
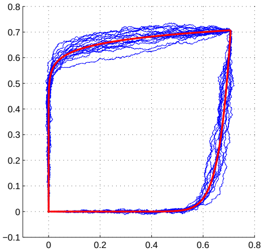


Figure: left panel: $\pi(x) \propto \alpha \exp\left(-(1/2)x'\Gamma_1^{-1}x\right) + (1-\alpha)\exp\left(-(1/2)x'\Gamma_2^{-1}x\right)$; right panel: Δ (green arrows) and Δ_∞ (blue arrows)



1.2
1
0.8
0.6
0.4
0.2
0
−0.2
−0.2
0
0.2
0.4
0.6
0.8
1
1.2



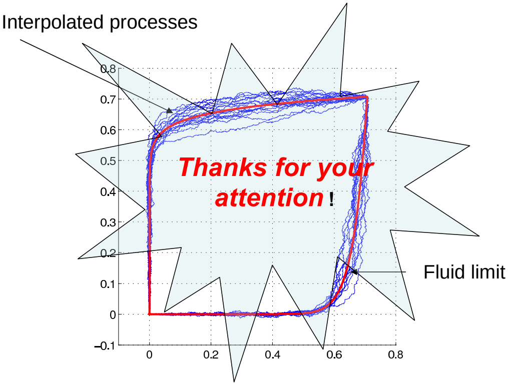
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
0
0.2
0.4
0.6
0.8

Conclusion

- ODE techniques provide a general and powerful approach to establishing stability and ergodic theorems for a Markov chain.
- In typical applications, the assumptions of this work hold for *any* $p > 0$ and consequently, the ergodic Theorem asserts that the mean of *any* function with polynomial growth converges to its steady-state mean faster than *any* polynomial rate. This result is optimal (under the stated assumptions) since it is impossible to obtain a geometric rate of convergence even when Δ, $\{\epsilon_k\}$ and the function f are bounded.

Research directions

1. The ODE method developed within the queueing networks research community has undergone many refinements, and has been applied in many very different contexts. Some of these extensions might serve well in MCMC
   - Design of control variates (to reduce Monte-Carlo variance or to design control methods)
   - Choice of sampling policies (analysis of the limiting ODE might help to understand the dynamic of the chain)
2. The ODE method may yield interesting results in situations where the Lyapunov functions are difficult to use (hybrid method, trans-dimensional MCMC - to do)
3. When $\pi(x)$ is polynomial in the tails, the renormalization should then be $\beta = 1$.. the fluid limits are then genuinely non-deterministic (but are diffusion) and all the previous theory need to be reconsidered (to do).
4. ....


Interpolated processes
Fluid limit
Thanks for your attention!
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
−0.1
0
0.2
0.4
0.6
0.8