Subgeometric hypocoercivity for piecewise-deterministic Markov process Monte Carlo methods

Paul Dobson

Joint work with C. Andrieu, A. Wang

Contents

Convergence to equilibrium

How fast does a process $\{Z_t\}_{t\geq 0}$ converge to equilibria?

We have a semigroup

$$\mathcal{P}_t f(z) = \mathbb{E}_z[f(Z_t)]$$

Target measure

$$\pi(dz) = \frac{e^{-U(z)}}{\mathcal{Z}} dz.$$

How fast does a process $\{Z_t\}_{t\geq 0}$ converge to equilibria?
We have a semigroup

$$\mathcal{P}_t f(z) = \mathbb{E}_z[f(Z_t)]$$

Target measure

$$\pi(dz) = \frac{e^{-U(z)}}{\mathcal{Z}} dz.$$

How fast does a process $\{Z_t\}_{t\geq 0}$ converge to equilibria?
We have a semigroup

$$\mathcal{P}_t f(z) = \mathbb{E}_z[f(Z_t)]$$

Target measure

$$\pi(dz) = \frac{e^{-U(z)}}{\mathcal{Z}} dz.$$

Convergence to equilibrium

How fast does a process $\{Z_t\}_{t\geq 0}$ converge to equilibria?
We have a semigroup

$$\mathcal{P}_t f(z) = \mathbb{E}_z[f(Z_t)]$$

Target measure

$$\pi(dz) = \frac{e^{-U(z)}}{\mathcal{Z}} dz.$$

Find λ such that

$$\|\mathcal{P}_t f - \pi(f)\|_{L^2_\pi} \leq e^{-\lambda t} \|f - \pi(f)\|_{L^2_\pi}.$$

Convergence to equilibrium

How fast does a process $\{Z_t\}_{t\geq 0}$ converge to equilibria?
We have a semigroup

$$\mathcal{P}_t f(z) = \mathbb{E}_z[f(Z_t)]$$

Target measure

$$\pi(dz) = \frac{e^{-U(z)}}{\mathcal{Z}} dz.$$

Find C, λ such that

$$\|\mathcal{P}_t f - \pi(f)\|_{L^2_\pi} \leq Ce^{-\lambda t}\|f - \pi(f)\|_{L^2_\pi}.$$

How fast does a process $\{Z_t\}_{t\geq 0}$ converge to equilibria?
We have a semigroup

$$\mathcal{P}_t f(z) = \mathbb{E}_z[f(Z_t)]$$

Target measure

$$\pi(dz) = \frac{e^{-U(z)}}{\mathcal{Z}} dz.$$

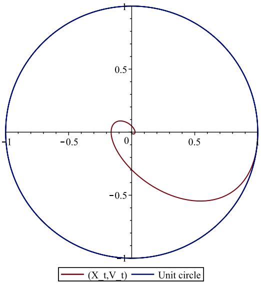
Find ξ such that

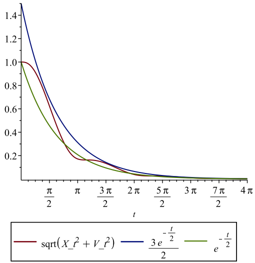
$$\|\mathcal{P}_t f - \pi(f)\|_{L^2_\pi} \leq \xi(t) \|f - \pi(f)\|_{L^2_\pi}.$$

Consider the ODE

$$\frac{d}{dt}\begin{pmatrix} X_t \\ V_t \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} X_t \\ V_t \end{pmatrix}$$

$$\frac{1}{2}\frac{d}{dt}|(X_t, V_t)|^2 = \begin{pmatrix} X_t & V_t \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} X_t \\ V_t \end{pmatrix} = -V_t^2.$$



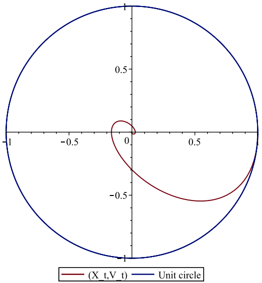


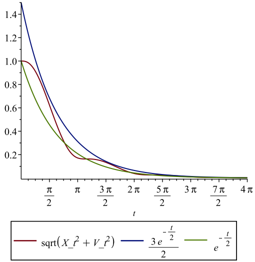
Basic Example

Consider the ODE

$$\frac{d}{dt}\begin{pmatrix} X_t \\ V_t \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} X_t \\ V_t \end{pmatrix}$$

$$\frac{1}{2}\frac{d}{dt}|(X_t, V_t)|^2 = \begin{pmatrix} X_t & V_t \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} X_t \\ V_t \end{pmatrix} = -V_t^2.$$



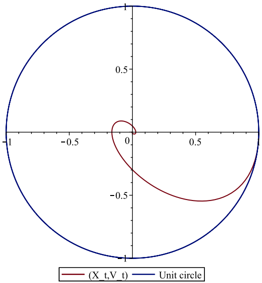


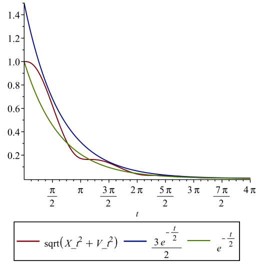
Consider the ODE

$$\frac{d}{dt}\begin{pmatrix} X_t \\ V_t \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} X_t \\ V_t \end{pmatrix}$$

For $(X_0, V_0) = (1, 0)$

$$\frac{1}{2}\frac{d}{dt}\bigg|_{t=0} |(X_t, V_t)|^2 = \begin{pmatrix} 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$





Strategy of Hypocoercivity

This is a strategy introduced by Dolbeault, Mouhot, and Schmeiser 2015. Let $\mathcal{L}$ be the generator corresponding to $\mathcal{P}_t$ and write

$$\mathcal{L} = S + T, \quad S^* = S, \quad T^* = -T.$$

Using $\partial_t \mathcal{P}_t f = \mathcal{L} f$ we write

$$\frac{d}{dt} \langle \mathcal{P}_t f, \mathcal{P}_t f \rangle_{L^2_\pi} = \langle \mathcal{L} \mathcal{P}_t f, \mathcal{P}_t f \rangle_{L^2_\pi} = \langle S \mathcal{P}_t f, \mathcal{P}_t f \rangle_{L^2_\pi}$$

If S is coercive,

$$\langle Sf, f \rangle_{L^2_\pi} \leq -\lambda \|f\|^2$$

then we can use Gronwall's inequality to show exponential convergence.
If S is hypocoercive,

$$\langle Sf, f \rangle_{L^2_\pi} \leq -\lambda \|(1 - \Pi) f\|^2$$

for some projection Π.

Strategy of Hypocoercivity

This is a strategy introduced by Dolbeault, Mouhot, and Schmeiser 2015. Let $\mathcal{L}$ be the generator corresponding to $\mathcal{P}_t$ and write

$$\mathcal{L} = S + T, \quad S^* = S, \quad T^* = -T.$$

Using $\partial_t \mathcal{P}_t f = \mathcal{L} f$ we write

$$\frac{d}{dt}\langle \mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle \mathcal{L}\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle S\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi}$$

If S is coercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|f\|^2$$

then we can use Gronwall's inequality to show exponential convergence.
If S is hypocoercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|(1-\Pi)f\|^2$$

for some projection Π.

Strategy of Hypocoercivity

This is a strategy introduced by Dolbeault, Mouhot, and Schmeiser 2015. Let $\mathcal{L}$ be the generator corresponding to $\mathcal{P}_t$ and write

$$\mathcal{L} = S + T, \quad S^* = S, \quad T^* = -T.$$

Using $\partial_t \mathcal{P}_t f = \mathcal{L} f$ we write

$$\frac{d}{dt}\langle \mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle \mathcal{L}\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle S\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi}$$

If S is coercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|f\|^2$$

then we can use Gronwall's inequality to show exponential convergence.
If S is hypocoercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|(1-\Pi)f\|^2$$

for some projection Π.

Strategy of Hypocoercivity

This is a strategy introduced by Dolbeault, Mouhot, and Schmeiser 2015. Let $\mathcal{L}$ be the generator corresponding to $\mathcal{P}_t$ and write

$$\mathcal{L} = S + T, \quad S^* = S, \quad T^* = -T.$$

Using $\partial_t \mathcal{P}_t f = \mathcal{L} f$ we write

$$\frac{d}{dt}\langle \mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle \mathcal{L}\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle S\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi}$$

If S is coercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|f\|^2$$

then we can use Gronwall's inequality to show exponential convergence.

If S is hypocoercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|(1-\Pi)f\|^2$$

for some projection Π.

Strategy of Hypocoercivity

This is a strategy introduced by Dolbeault, Mouhot, and Schmeiser 2015. Let $\mathcal{L}$ be the generator corresponding to $\mathcal{P}_t$ and write

$$\mathcal{L} = S + T, \quad S^* = S, \quad T^* = -T.$$

Using $\partial_t \mathcal{P}_t f = \mathcal{L} f$ we write

$$\frac{d}{dt}\langle \mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle \mathcal{L}\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi} = \langle S\mathcal{P}_t f, \mathcal{P}_t f\rangle_{L^2_\pi}$$

If S is coercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|f\|^2$$

then we can use Gronwall's inequality to show exponential convergence.
If S is hypocoercive,

$$\langle Sf, f\rangle_{L^2_\pi} \leq -\lambda \|(1-\Pi)f\|^2$$

for some projection Π.

Strategy of Hypocoercivity

$$\mathscr{H}(f) = \frac{1}{2}\|f\|^2_{L^2_\pi} + \varepsilon\langle Af, f\rangle_{L^2_\pi}.$$

Then find conditions to ensure

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) \leq -\lambda\mathscr{H}(\mathcal{P}_t f).$$

We can expand the LHS

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \langle Sf, f\rangle - \varepsilon\langle AT\Pi f, f\rangle + \varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

If we choose $A = (1+(T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and $\phi(s) = \frac{s}{1+s}$

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \underbrace{\varepsilon\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$
$$+\varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Strategy of Hypocoercivity

$$\mathscr{H}(f) = \frac{1}{2}\|f\|^2_{L^2_\pi} + \varepsilon\langle Af, f\rangle_{L^2_\pi}.$$

Then find conditions to ensure

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) \leq -\lambda \mathscr{H}(\mathcal{P}_t f).$$

We can expand the LHS

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \langle Sf, f\rangle - \varepsilon\langle AT\Pi f, f\rangle + \varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

If we choose $A = (1 + (T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and $\phi(s) = \frac{s}{1+s}$

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} \underbrace{- \varepsilon\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$

$$+\varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Strategy of Hypocoercivity

$$\mathscr{H}(f) = \frac{1}{2}\|f\|^2_{L^2_\pi} + \varepsilon\langle Af, f\rangle_{L^2_\pi}.$$

Then find conditions to ensure

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) \le -\lambda \mathscr{H}(\mathcal{P}_t f).$$

We can expand the LHS

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \langle Sf, f\rangle - \varepsilon\langle AT\Pi f, f\rangle + \varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

If we choose $A = (1 + (T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and $\phi(s) = \frac{s}{1+s}$

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} \underbrace{-\varepsilon\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$

$$+\varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Strategy of Hypocoercivity

$$\mathscr{H}(f) = \frac{1}{2}\|f\|^2_{L^2_\pi} + \varepsilon\langle Af, f\rangle_{L^2_\pi}.$$

Then find conditions to ensure

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) \leq -\lambda \mathscr{H}(\mathcal{P}_t f).$$

We can expand the LHS

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \langle Sf, f\rangle - \varepsilon\langle AT\Pi f, f\rangle + \varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

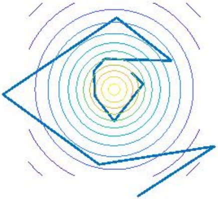
If we choose $A = (1 + (T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and $\phi(s) = \frac{s}{1+s}$

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \underbrace{\varepsilon\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$
$$+\varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Piecewise Deterministic Markov processes $\{(X_t, V_t)\}_{t\geq 0}$ have three components:

- Deterministic motion, for this talk before any events:

$$(x, v) \mapsto (x + tv, v)$$



- Event times, T_i random times which occur at a rate $\lambda(x, v)$, that is

$$\mathbb{P}(T_1 > t) = \exp\left(-\int_0^t \lambda(X_s, V_s)ds\right).$$

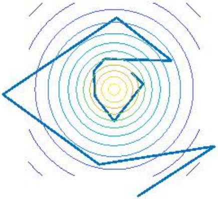
- Events, we shall consider two types of events reflections and refreshments.

We extend the target measure $\pi(dx)$ to $\mu(dx, dv) = \pi(dx)\nu(dv)$.

Piecewise Deterministic Markov processes $\{(X_t, V_t)\}_{t\geq 0}$ have three components:

- Deterministic motion, for this talk before any events:

$$(x, v) \mapsto (x + tv, v)$$



- Event times, T_i random times which occur at a rate $\lambda(x, v)$, that is

$$\mathbb{P}(T_1 > t) = \exp\left(-\int_0^t \lambda(X_s, V_s)ds\right).$$

- Events, we shall consider two types of events reflections and refreshments.

We extend the target measure $\pi(dx)$ to $\mu(dx, dv) = \pi(dx)\nu(dv)$.

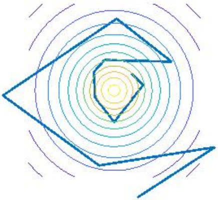
Piecewise Deterministic Markov processes $\{(X_t, V_t)\}_{t\geq 0}$ have three components:

- Deterministic motion, for this talk before any events:

$$(x, v) \mapsto (x + tv, v)$$

- Event times, T_i random times which occur at a rate $\lambda(x, v)$, that is

$$\mathbb{P}(T_1 > t) = \exp\left(-\int_0^t \lambda(X_s, V_s)ds\right).$$



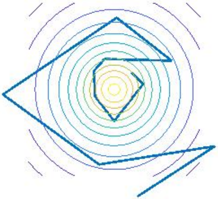
- Events, we shall consider two types of events reflections and refreshments.

We extend the target measure $\pi(dx)$ to $\mu(dx, dv) = \pi(dx)\nu(dv)$.

Piecewise Deterministic Markov processes $\{(X_t, V_t)\}_{t\geq 0}$ have three components:

- Deterministic motion, for this talk before any events:

$$(x, v) \mapsto (x + tv, v)$$



- Event times, T_i random times which occur at a rate $\lambda(x, v)$, that is

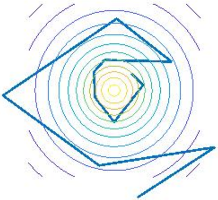
$$\mathbb{P}(T_1 > t) = \exp\left(-\int_0^t \lambda(X_s, V_s)ds\right).$$

- Events, we shall consider two types of events reflections and refreshments.

We extend the target measure $\pi(dx)$ to $\mu(dx, dv) = \pi(dx)\nu(dv)$.

Piecewise Deterministic Markov processes $\{(X_t, V_t)\}_{t \geq 0}$ have three components:

- Deterministic motion, for this talk before any events:
$$(x, v) \mapsto (x + tv, v)$$



- Event times, T_i random times which occur at a rate $\lambda(x, v)$, that is
$$\mathbb{P}(T_1 > t) = \exp\left(-\int_0^t \lambda(X_s, V_s) ds\right).$$

- Events, we shall consider two types of events reflections and refreshments.

We extend the target measure $\pi(dx)$ to $\mu(dx, dv) = \pi(dx)\nu(dv)$.

PDMP and examples

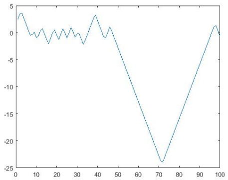
We will summarise PDMP by its generator

$$\mathcal{L}f(x,v) = \underbrace{v^T \nabla_x f(x,v)}_{\text{deterministic}} + \underbrace{\sum_{k=1}^{K} \lambda_k(x,v)[(\mathcal{B}_k - 1)f](x,v)}_{\text{reflections}} + \underbrace{\lambda_{\mathrm{r}}(\Pi - 1)f(x,v)}_{\text{refreshments}}.$$

$$\Pi f(x,v) = \int f(x,w)\nu(dw).$$

We ensure this has invariant measure $\mu(dx,dv) = \mathcal{Z}^{-1}e^{-U(x)}dx\nu(dv)$.

- Zig Zag: $K = d$, $\mathcal{B}_k f(x,v) = f(x, v_1, \ldots, -v_k, \ldots, v_d)$, and $\lambda_k(x,v) = (v_k \partial_{x_k} U(x))^+$.



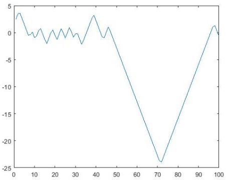
We will summarise PDMP by its generator

$$\mathcal{L}f(x,v) = \underbrace{v^T \nabla_x f(x,v)}_{\text{deterministic}} + \underbrace{\sum_{k=1}^{K} \lambda_k(x,v)[(\mathcal{B}_k - 1)f](x,v)}_{\text{reflections}} + \underbrace{\lambda_{\mathrm{r}}(\Pi - 1)f(x,v)}_{\text{refreshments}}.$$

$$\Pi f(x,v) = \int f(x,w)\nu(dw).$$

We ensure this has invariant measure $\mu(dx,dv) = \mathcal{Z}^{-1} e^{-U(x)} dx \nu(dv)$.

- Zig Zag: $K = d$, $\mathcal{B}_k f(x,v) = f(x, v_1, \ldots, -v_k, \ldots, v_d)$, and $\lambda_k(x,v) = (v_k \partial_{x_k} U(x))^+$.



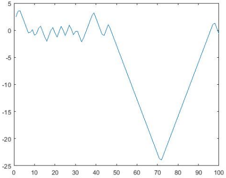
We will summarise PDMP by its generator

$$\mathcal{L}f(x,v) = \underbrace{v^T \nabla_x f(x,v)}_{\text{deterministic}} + \underbrace{\sum_{k=1}^{K} \lambda_k(x,v)[(\mathcal{B}_k - 1)f](x,v)}_{\text{reflections}} + \underbrace{\lambda_{\mathrm{r}}(\Pi - 1)f(x,v)}_{\text{refreshments}}.$$

$$\Pi f(x,v) = \int f(x,w)\nu(dw).$$

We ensure this has invariant measure $\mu(dx,dv) = \mathcal{Z}^{-1} e^{-U(x)} dx \nu(dv)$.

- Zig Zag: $K = d$,
  $\mathcal{B}_k f(x,v) = f(x, v_1, \ldots, -v_k, \ldots, v_d)$, and
  $\lambda_k(x,v) = (v_k \partial_{x_k} U(x))^+$.



Bouncy Particle Sampler

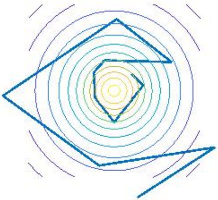
For the purpose of this talk we will only talk about the Bouncy Particle Sampler.

$$\mathcal{L}f(x,v) = \underbrace{v^T\nabla_x f(x,v)}_{\text{deterministic}} + \underbrace{(v\cdot\nabla_x U(x))^+[f(x,R(x)v) - f(x,v)]}_{\text{reflections}} + \underbrace{\lambda_{\mathrm{r}}(\Pi - 1)f(x,v)}_{\text{refreshments}}.$$

Here ν is a standard Gaussian measure and

$$\Pi f(x,v) = \int f(x,w)\nu(dw),$$

$$R(x)v = v - 2\frac{v\cdot\nabla_x U(x)}{|\nabla_x U(x)|^2}\nabla_x U(x).$$



The geometric case was considered by Andrieu, Durmus, Nüsken, and Roussel 2018.

$$\mathcal{L}f(x,v) = v^T\nabla_x f(x,v) + (v\cdot\nabla_x U(x))^+[f(x,R(x)v) - f(x,v)] + \lambda_{\mathrm{r}}(\Pi - 1)f(x,v)$$

we decompose this as

$$Sf(x,v) = \lambda^e(x,v)[f(x,R(x)v) - f(x,v)] + \lambda_{\mathrm{r}}(\Pi - 1)f(x,v)$$
$$Tf(x,v) = v^T\nabla_x f(x,v) + \lambda^o(x,v)[f(x,R(x)v) - f(x,v)].$$

Here

$$\lambda^e(x,v) = \frac{1}{2}(\lambda(x,v) + \lambda(x,-v)) = |v^T\nabla_x U(x)|,$$
$$\lambda^o(x,v) = \frac{1}{2}(\lambda(x,v) - \lambda(x,-v)) = v^T\nabla_x U(x).$$

Hypocoercivity for PDMP: Geometric Case

The geometric case was considered by Andrieu, Durmus, Nüsken, and Roussel 2018.

$$\mathcal{L}f(x,v) = v^T\nabla_x f(x,v) + (v\cdot\nabla_x U(x))^+[f(x,R(x)v) - f(x,v)] \\ + \lambda_{\mathrm{r}}(\Pi - 1)f(x,v)$$

we decompose this as

$$Sf(x,v) = \lambda^e(x,v)[f(x,R(x)v) - f(x,v)] + \lambda_{\mathrm{r}}(\Pi - 1)f(x,v)$$
$$Tf(x,v) = v^T\nabla_x f(x,v) + \lambda^o(x,v)[f(x,R(x)v) - f(x,v)].$$

Here

$$\lambda^e(x,v) = \frac{1}{2}(\lambda(x,v) + \lambda(x,-v)) = |v^T\nabla_x U(x)|,$$
$$\lambda^o(x,v) = \frac{1}{2}(\lambda(x,v) - \lambda(x,-v)) = v^T\nabla_x U(x).$$

Hypocoercivity for PDMP: Geometric Case

The geometric case was considered by Andrieu, Durmus, Nüsken, and Roussel 2018.

$$\mathcal{L}f(x,v) = v^T\nabla_x f(x,v) + (v\cdot\nabla_x U(x))^+[f(x,R(x)v) - f(x,v)] \\ + \lambda_{\mathrm{r}}(\Pi - 1)f(x,v)$$

we decompose this as

$$Sf(x,v) = \lambda^e(x,v)[f(x,R(x)v) - f(x,v)] + \lambda_{\mathrm{r}}(\Pi - 1)f(x,v)$$
$$Tf(x,v) = v^T\nabla_x f(x,v) + \lambda^o(x,v)[f(x,R(x)v) - f(x,v)].$$

Here

$$\lambda^e(x,v) = \frac{1}{2}(\lambda(x,v) + \lambda(x,-v)) = |v^T\nabla_x U(x)|,$$
$$\lambda^o(x,v) = \frac{1}{2}(\lambda(x,v) - \lambda(x,-v)) = v^T\nabla_x U(x).$$

Recall we highlighted two terms

$$\underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \underbrace{\varepsilon\langle\phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$

For PDMP microscopic coercivity is immediate:

$$\langle Sf, f\rangle \leq -\lambda_{\mathrm{r}}\|(1-\Pi)f\|^2$$

The operator $(T\Pi)^*(T\Pi)$ has a simple form

$$-(T\Pi)^*(T\Pi)f = \Delta\Pi f - (\nabla U(x))^T\nabla\Pi f =: \mathcal{A}\Pi f.$$

Macroscopic coercivity follows provided we have a Poincaré Inequality for $\mathcal{A}$, i.e.

$$\langle\mathcal{A}g, g\rangle \leq -C_P\|g\|^2.$$

The subgeometric case is when the Poincaré inequality does not hold. This case was considered for SDEs by Grothaus and Wang 2019.

Recall we highlighted two terms

$$\underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} \quad -\underbrace{\varepsilon\langle\phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$

For PDMP microscopic coercivity is immediate:

$$\langle Sf, f\rangle \leq -\lambda_{\mathrm{r}}\|(1-\Pi)f\|^2$$

The operator $(T\Pi)^*(T\Pi)$ has a simple form

$$-(T\Pi)^*(T\Pi)f = \Delta\Pi f - (\nabla U(x))^T\nabla\Pi f =: \mathcal{A}\Pi f.$$

Macroscopic coercivity follows provided we have a Poincaré Inequality for $\mathcal{A}$, i.e.

$$\langle\mathcal{A}g, g\rangle \leq -C_P\|g\|^2.$$

The subgeometric case is when the Poincaré inequality does not hold. This case was considered for SDEs by Grothaus and Wang 2019.

Recall we highlighted two terms

$$\underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \underbrace{\varepsilon\langle\phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$

For PDMP microscopic coercivity is immediate:

$$\langle Sf, f\rangle \leq -\lambda_{\mathrm{r}}\|(1-\Pi)f\|^2$$

The operator $(T\Pi)^*(T\Pi)$ has a simple form

$$-(T\Pi)^*(T\Pi)f = \Delta\Pi f - (\nabla U(x))^T\nabla\Pi f =: \mathcal{A}\Pi f.$$

Macroscopic coercivity follows provided we have a Poincaré Inequality for $\mathcal{A}$, i.e.

$$\langle\mathcal{A}g, g\rangle \leq -C_P\|g\|^2.$$

The subgeometric case is when the Poincaré inequality does not hold. This case was considered for SDEs by Grothaus and Wang 2019.

Recall we highlighted two terms

$$\underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} \quad - \varepsilon\underbrace{\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$

For PDMP microscopic coercivity is immediate:

$$\langle Sf, f\rangle \leq -\lambda_{\mathrm{r}}\|(1-\Pi)f\|^2$$

The operator $(T\Pi)^*(T\Pi)$ has a simple form

$$-(T\Pi)^*(T\Pi)f = \Delta\Pi f - (\nabla U(x))^T\nabla\Pi f =: \mathcal{A}\Pi f.$$

Macroscopic coercivity follows provided we have a Poincaré Inequality for $\mathcal{A}$, i.e.

$$\langle \mathcal{A}g, g\rangle \leq -C_P\|g\|^2.$$

The subgeometric case is when the Poincaré inequality does not hold. This case was considered for SDEs by Grothaus and Wang 2019.

Recall we highlighted two terms

$$\underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} \quad -\varepsilon\underbrace{\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$

For PDMP microscopic coercivity is immediate:

$$\langle Sf, f\rangle \leq -\lambda_{\mathrm{r}}\|(1-\Pi)f\|^2$$

The operator $(T\Pi)^*(T\Pi)$ has a simple form

$$-(T\Pi)^*(T\Pi)f = \Delta\Pi f - (\nabla U(x))^T\nabla\Pi f =: \mathcal{A}\Pi f.$$

Macroscopic coercivity follows provided we have a Poincaré Inequality for $\mathcal{A}$, i.e.

$$\langle \mathcal{A}g, g\rangle \leq -C_P\|g\|^2.$$

The subgeometric case is when the Poincaré inequality does not hold. This case was considered for SDEs by Grothaus and Wang 2019.

Weak Poincaré Inequality

$\mathcal{A}$ satisfies the Poincaré Inequality if

$$\|g\|^2 \leq -C_P^{-1}\langle \mathcal{A}g, g\rangle.$$

$\mathcal{A}$ satisfies the weak Poincaré Inequality if for all $r > 0$

$$\|g\|^2 \leq -\alpha(r)\langle \mathcal{A}g, g\rangle + r\,\mathrm{osc}(g)^2.$$

Here $\mathrm{osc}(g) = \sup g - \inf g$.
Unlike the Poincaré inequality, the weak Poincaré inequality is generally satisfied by any operator of the form $\mathcal{A} = \Delta + \nabla U^T \nabla$ as shown by Röckner and Wang 2001.
However $\alpha(r)$ can be difficult determine precisely.

Weak Poincaré Inequality

$\mathcal{A}$ satisfies the Poincaré Inequality if

$$\|g\|^2 \leq -C_P^{-1}\langle \mathcal{A}g, g\rangle.$$

$\mathcal{A}$ satisfies the weak Poincaré Inequality if for all $r > 0$

$$\|g\|^2 \leq -\alpha(r)\langle \mathcal{A}g, g\rangle + r\,\mathrm{osc}(g)^2.$$

Here $\mathrm{osc}(g) = \sup g - \inf g$.

Unlike the Poincaré inequality, the weak Poincaré inequality is generally satisfied by any operator of the form $\mathcal{A} = \Delta + \nabla U^T \nabla$ as shown by Röckner and Wang 2001.

However $\alpha(r)$ can be difficult determine precisely.

$\mathcal{A}$ satisfies the Poincaré Inequality if

$$\|g\|^2 \leq -C_P^{-1}\langle \mathcal{A}g, g\rangle.$$

$\mathcal{A}$ satisfies the weak Poincaré Inequality if for all $r > 0$

$$\|g\|^2 \leq -\alpha(r)\langle \mathcal{A}g, g\rangle + r\mathrm{osc}(g)^2.$$

Here $\mathrm{osc}(g) = \sup g - \inf g$.
Unlike the Poincaré inequality, the weak Poincaré inequality is generally satisfied by any operator of the form $\mathcal{A} = \Delta + \nabla U^T \nabla$ as shown by Röckner and Wang 2001.
However $\alpha(r)$ can be difficult determine precisely.

$\mathcal{A}$ satisfies the Poincaré Inequality if

$$\|g\|^2 \leq -C_P^{-1}\langle \mathcal{A}g, g\rangle.$$

$\mathcal{A}$ satisfies the weak Poincaré Inequality if for all $r > 0$

$$\|g\|^2 \leq -\alpha(r)\langle \mathcal{A}g, g\rangle + r\,\mathrm{osc}(g)^2.$$

Here $\mathrm{osc}(g) = \sup g - \inf g$.
Unlike the Poincaré inequality, the weak Poincaré inequality is generally satisfied by any operator of the form $\mathcal{A} = \Delta + \nabla U^T \nabla$ as shown by Röckner and Wang 2001.
However $\alpha(r)$ can be difficult determine precisely.

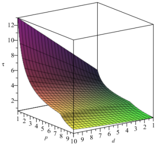
The weak Poincaré Inequality follows from a local Poincaré inequality: For any $\varepsilon > 0$ there exists $A \subseteq \mathbb{R}^N$, $c > 0$ s.t. $\mu(A) \geq 1 - \varepsilon$ and

$$\mu(g\mathbb{1}_A) \leq -c\langle \mathcal{A}g, g\rangle + \frac{\mu(g\mathbb{1}_A)^2}{\mu(A)}.$$

Examples:

- For $U(x) = -(d+p)\log(1+|x|)$ we have $\alpha(r) = c(1 + r^{-\tau})$,

$$\tau = \min\left(\frac{d+p+2}{p}, \frac{4p+4+2d}{(p^2-4-2d-2p)^+}\right).$$



- For $U(x) = -\sigma|x|^{\delta}$, with $\delta \in (0,1)$, then

$$\alpha(r) = c[1 + \log(1 + r^{-1})]^{4(1-\delta)/\delta}$$

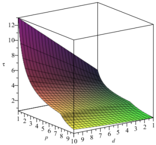
The weak Poincaré Inequality follows from a local Poincaré inequality: For any $\varepsilon > 0$ there exists $A \subseteq \mathbb{R}^N$, $c > 0$ s.t. $\mu(A) \geq 1 - \varepsilon$ and

$$\mu(g\mathbb{1}_A) \leq -c\langle \mathcal{A}g, g\rangle + \frac{\mu(g\mathbb{1}_A)^2}{\mu(A)}.$$

Examples:

- For $U(x) = -(d+p)\log(1+|x|)$ we have $\alpha(r) = c(1 + r^{-\tau})$,

$$\tau = \min\left(\frac{d+p+2}{p}, \frac{4p+4+2d}{(p^2-4-2d-2p)^+}\right).$$



- For $U(x) = -\sigma|x|^\delta$, with $\delta \in (0,1)$, then

$$\alpha(r) = c[1 + \log(1 + r^{-1})]^{4(1-\delta)/\delta}$$

Weak Poincaré Inequality

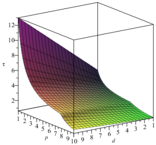
The weak Poincaré Inequality follows from a local Poincaré inequality: For any $\varepsilon > 0$ there exists $A \subseteq \mathbb{R}^N$, $c > 0$ s.t. $\mu(A) \geq 1 - \varepsilon$ and

$$\mu(g\mathbb{1}_A) \leq -c\langle \mathcal{A}g, g\rangle + \frac{\mu(g\mathbb{1}_A)^2}{\mu(A)}.$$

Examples:

- For $U(x) = -(d+p)\log(1+|x|)$ we have $\alpha(r) = c(1 + r^{-\tau})$,

$$\tau = \min\left(\frac{d+p+2}{p}, \frac{4p+4+2d}{(p^2-4-2d-2p)^+}\right).$$



- For $U(x) = -\sigma|x|^\delta$, with $\delta \in (0,1)$, then

$$\alpha(r) = c[1 + \log(1 + r^{-1})]^{4(1-\delta)/\delta}$$

Remaining terms

Recall $A = (1 + (T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \underbrace{\varepsilon\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$
$$+\varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Note TA is a bounded operator. The remaining terms we can rewrite as

$$|\langle ASf, f\rangle| \leq \|(1-\Pi)f\|(\lambda_{\mathrm{r}}\|\nabla_x u_f\| + 2\sqrt{2}\sum_{k=1}^{K}\|\nabla_x U(x)^T\nabla_x u_f\|).$$

There is a similar expression for $AT(1-\Pi)$. Here u_f is the solution of

$$(1 + (T\Pi)^*(T\Pi))u_f = \quad (1 + \nabla_x^*\nabla_x)u_f = \Pi f.$$

Remaining terms

Recall $A = (1 + (T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \varepsilon\underbrace{\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$
$$+\varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Note TA is a bounded operator. The remaining terms we can rewrite as

$$|\langle ASf, f\rangle| \leq \|(1-\Pi)f\|(\lambda_{\mathrm{r}}\|\nabla_x u_f\| + 2\sqrt{2}\sum_{k=1}^{K}\|\nabla_x U(x)^T\nabla_x u_f\|).$$

There is a similar expression for $AT(1-\Pi)$. Here u_f is the solution of

$$(1 + (T\Pi)^*(T\Pi))u_f = (1 + \nabla_x^*\nabla_x)u_f = \Pi f.$$

Remaining terms

Recall $A = (1 + (T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \underbrace{\varepsilon\langle \phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}}$$
$$+\varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Note TA is a bounded operator. The remaining terms we can rewrite as

$$|\langle ASf, f\rangle| \le \|(1-\Pi)f\|(\lambda_{\mathrm{r}}\|\nabla_x u_f\| + 2\sqrt{2}\sum_{k=1}^{K}\|\nabla_x U(x)^T\nabla_x u_f\|).$$

There is a similar expression for $AT(1-\Pi)$. Here u_f is the solution of

$$(1 + (T\Pi)^*(T\Pi))u_f = \quad (1 + \nabla_x^*\nabla_x)u_f = \Pi f.$$

Remaining terms

Recall $A = (1 + (T\Pi)^*(T\Pi))^{-1}(-T\Pi)^*$ and

$$\frac{d}{dt}\mathscr{H}(\mathcal{P}_t f) = \underbrace{\langle Sf, f\rangle}_{\text{Microscopic Coercivity}} - \underbrace{\varepsilon\langle\phi((T\Pi)^*(T\Pi))f, f\rangle}_{\text{Macroscopic Coercivity}} + \varepsilon\langle ASf, f\rangle - \varepsilon\langle AT(I-\Pi)f, f\rangle + \varepsilon\langle TAf, f\rangle$$

Note TA is a bounded operator. The remaining terms we can rewrite as

$$|\langle ASf, f\rangle| \leq \|(1-\Pi)f\|(\lambda_{\mathrm{r}}\|\nabla_x u_f\| + 2\sqrt{2}\sum_{k=1}^{K}\|\nabla_x U(x)^T\nabla_x u_f\|).$$

There is a similar expression for $AT(1-\Pi)$. Here u_f is the solution of

$$(1 + (T\Pi)^*(T\Pi))u_f = \quad (1 + \nabla_x^*\nabla_x)u_f = \Pi f.$$

Poisson Equation estimates

For

$$(1 + \nabla_x^* \nabla_x) u_f = \Pi f.$$

Then there exists $\kappa_1, \kappa_2 > 0$ such that

$$\begin{aligned} \max\{\|u_f\|, \|\nabla_x u_f\|\} &\leq \|\Pi f\| \\ \|\nabla_x^2 u_f\| &\leq \kappa_1 \|\Pi f\| \\ \|(\nabla_x U)^T \nabla_x u_f\| &\leq \kappa_2 \|\Pi f\| \end{aligned}$$

This was shown by Lorenzi and Lunardi 2006.

Theorem

We have convergence of the semigroup for f with $\mu(f) = 0$

$$\|\mathcal{P}_t f\|^2 \leq \xi(t) \left(\|f\|^2 + \mathrm{osc}(f)^2 \right), \quad \forall t \geq 0$$

where

$$\xi(t) := c_1 \inf\{r > 0 : c_2 t \geq \alpha(r)^2 \log(1/r)\}.$$

Examples

Recall the two examples:

- For $U(x) = -(d+p)\log(1+|x|)$ we have $\alpha(r) = c(1+r^{-\tau})$,

$$\tau = \min\left(\frac{d+p+2}{p}, \frac{4p+4+2d}{(p^2-4-2d-2p)^+}\right).$$

Then

$$\xi(t) = t^{-\frac{1}{2\tau}}.$$

- For $U(x) = -\sigma|x|^{\delta}$, with $\delta \in (0,1)$, then

$$\alpha(r) = c[1+\log(1+r^{-1})]^{4(1-\delta)/\delta}$$

Then there exists k

$$\xi(t) = \exp\left(-kt^{\frac{\delta}{8-7\delta}}\right)$$

Examples

Recall the two examples:

- For $U(x) = -(d+p)\log(1+|x|)$ we have $\alpha(r) = c(1+r^{-\tau})$,

$$\tau = \min\left(\frac{d+p+2}{p}, \frac{4p+4+2d}{(p^2-4-2d-2p)^+}\right).$$

  Then

$$\xi(t) = t^{-\frac{1}{2\tau}}.$$

- For $U(x) = -\sigma|x|^\delta$, with $\delta \in (0,1)$, then

$$\alpha(r) = c[1+\log(1+r^{-1})]^{4(1-\delta)/\delta}$$

  Then there exists k

$$\xi(t) = \exp\left(-kt^{\frac{\delta}{8-7\delta}}\right)$$

Andrieu, Christophe, Paul Dobson, and Andi Q Wang (2020). "Subgeometric hypocoercivity for piecewise-deterministic Markov process Monte Carlo methods". In: *arXiv preprint arXiv:2011.09341*.

Andrieu, Christophe, Alain Durmus, Nikolas Nüsken, and Julien Roussel (2018). "Hypocoercivity of Piecewise Deterministic Markov Process-Monte Carlo". In: arXiv: `1808.08592`.

Dolbeault, Jean, Clément Mouhot, and Christian Schmeiser (2015). "Hypocoercivity for linear kinetic equations conserving mass". In: *Transactions of the American Mathematical Society* 367.6, pp. 6012–6019.

Grothaus, Martin and Feng-Yu Wang (2019). "Weak Poincaré inequalities for convergence rate of degenerate diffusion processes". In: *Annals of Probability* 47.5, pp. 2930–2952. ISSN: 2168894X. DOI: `10.1214/18-AOP1328`.

Lorenzi, Luca and Alessandra Lunardi (2006). "Elliptic operators with unbounded diffusion coefficients in L^2 spaces with respect to invariant measures". In: *Journal of Evolution Equations* 6, pp. 691–709. DOI: `10.1007/s00028-006-0283-6`.

Röckner, Michael and Feng-Yu Wang (2001). "Weak Poincaré Inequalities and L^2 Convergence Rates of Markov Semigroups". In: *Journal of Functional Analysis* 185, pp. 564–603.