Spectral theory and model reduction for complex Markov processes

Sean Meyn

Department of Electrical and Computer Engineering
University of Illinois
and the Coordinated Science Laboratory

Joint work with W. Huisinga, I. Kontoyiannis, and C. Schuette
Supported in part by the National Science Founation, ECS

Outline

Outline

Generator for an ODE

$$\frac{d}{dt}x(t) = \delta(x(t))$$

The generator is defined for C^1 functions ,

$$\mathcal{D}h\,(x) := \frac{d}{dt}h(x(t))\Big|_{\substack{t=0\\x=x(0)}} = \langle \delta(x), \nabla h(x) \rangle$$

Fundamental Theorem of Calculus,

$$h(x(T)) = h(x(0)) + \int_0^T \mathcal{D}h\,(x(t))\,dt$$

Generator for an SDE $dX(t) = \delta(X(t))\,dt + B(X(t))dN(t)$

The generator is defined for C^2 (extra nice) functions ,

$$\mathcal{D}h\,(x) := \frac{d}{dt}\mathsf{E}[h(X(t))\ |X(0)=x]\Big|_{t=0}$$

$$= \langle \delta(x), \nabla h(x)\rangle + \tfrac{1}{2}\mathrm{trace}\,(\Sigma(x)\nabla^2 h(x))$$

$$\Sigma(x) = B(x)B(x)^T$$

Martingale Representation

$$\mathsf{E}\Big[h(X(T)) - \int_0^T \mathcal{D}h\,(X(t))\,dt\Big] = h(X(0))$$

Eigenfunction for an ODE $\mathcal{D}h = \lambda h$

$z(t) = h(x(t))$ solution to linear ODE,

$$\frac{d}{dt} z(t) = \lambda z(t)$$

Applications

Decomposition:

If λ is nearly zero, level sets of h are *almost invariant*

The set $\{\, h(x) > 0 \,\}$ *is invariant*

Eigenfunction for an ODE $\mathcal{D}h = \lambda h$

$z(t) = h(x(t))$ solution to linear ODE,

$$\frac{d}{dt} z(t) = \lambda z(t)$$

Applications

Model reduction: $z(t) = (z_1(t), \ldots, z_n(t)) = (h_1(x(t)), \ldots, h_n(x(t)))$

Simple linear description,

$$\frac{d}{dt} z(t) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} z(t)$$

Eigenfunction for an SDE $\mathcal{D}h = \lambda h$

$Z(t) = h(X(t))$ solution to *almost* linear SDE,

$$dZ(t) = \lambda Z(t) + dN_Z(t)$$

white noise

$$dN_Z(t) := \nabla h\,(X(t)) \cdot B(X(t))dN_X(t)$$

Applications ...

Decompositions ? model reduction ??

This is the subject of this talk!

Outline

Exponential Ergodicity

Condition (V4): For some $V : \mathsf{X} \to [1, \infty)$, $b < \infty$, $\delta > 0$

K compact

$$\mathcal{D}V \leq -\delta V + b\mathbb{I}_K$$

Under a density assumption, the diffusion is exponentially ergodic,

There is a spectral gap

Exponential Ergodicity

Condition (V4): For some $V\colon \mathsf{X} \to [1,\infty)$, $b < \infty$, $\delta > 0$
K compact

$$\mathcal{D}V \le -\delta V + b\mathbb{I}_K$$

Example:

$$dX(t) = AX(t)dt + BdN(t)$$

(A,B) controllable
$\lambda(A) < 0$

$$V(x) = 1 + \tfrac{1}{2}\varepsilon x^T M x \qquad M > 0 \qquad A^T M + MA = -I$$

Exponential Ergodicity

Condition (V4): For some $V\colon \mathsf{X} \to [1,\infty)$, $b<\infty$, $\delta>0$
K compact

$$\mathcal{D}V \le -\delta V + b\mathbb{I}_K$$

Exponential ergodicity?

Choose weighted norm, $L_\infty^V = \left\{ f : \|f\|_V \triangleq \sup_{x\in\mathsf{X}} \frac{|f(x)|}{V(x)} < \infty \right\}$

$$|||P^t - \pi|||_V \triangleq \sup_{\|f\|_V=1} \Big(\|P^t f - \pi(f)\|_V \Big) \to 0$$ *Exponentially fast*

Down, M., Tweedie '95
Kontoyiannis & M. '03

Exponential Ergodicity

Condition (V4): For some $V\colon \mathsf{X} \to [1,\infty)$, $b<\infty$, $\delta>0$
K compact

$$\mathcal{D}V \le -\delta V + b\mathbb{I}_K$$

Spectral gap?

Choose weighted norm, $L_\infty^V = \Big\{ f : \|f\|_V \triangleq \sup_{x\in\mathsf{X}} \frac{|f(x)|}{V(x)} < \infty \Big\}$

$[Iz - \mathcal{D}]^{-1}$ Isolated pole at $z=0$ $\qquad \mathcal{D}\mathbf{1} = 0 \quad \pi\mathcal{D} = 0$

$[Iz - (\mathcal{D} - \mathbf{1}\otimes\pi)]^{-1}$ exists for $z \sim 0$

Down, M., Tweedie '95
Kontoyiannis & M. '03

Drift condition of Donsker & Varadhan

The nonlinear generator is defined by, $\mathcal{H}(F) = e^{-F}\mathcal{D}e^{F}$

Condition (DV3): For some $V, W : \mathsf{X} \to [1,\infty),\ b < \infty,$

W has compact sublevel sets

$$\mathcal{H}(V) \le -W + b$$

Drift condition of Donsker & Varadhan

The nonlinear generator is defined by, $\mathcal{H}(F) = e^{-F}\mathcal{D}e^{F}$

Condition (DV3): For some $V, W : \mathsf{X} \to [1,\infty),\ b < \infty,$

W has compact sublevel sets

$$\mathcal{H}(V) \le -W + b$$

Example: $dX(t) = AX(t)dt + BdN(t)$ (A, B) controllable $\lambda(A) < 0$

$$V(x) = 1 + \tfrac{1}{2}\varepsilon x^T M x \qquad M > 0 \quad A^T M + MA = -I$$

$$W(x) = 1 + \delta\|x\|$$

Drift condition of Donsker & Varadhan

The nonlinear generator is defined by, $\mathcal{H}(F) = e^{-F}\mathcal{D}e^{F}$

Condition (DV3): For some $V, W : \mathsf{X} \to [1,\infty),\ b < \infty,$

W has compact sublevel sets

$$\mathcal{H}(V) \leq -W + b$$

This is a relaxation of D&V's conditions for Large Deviations

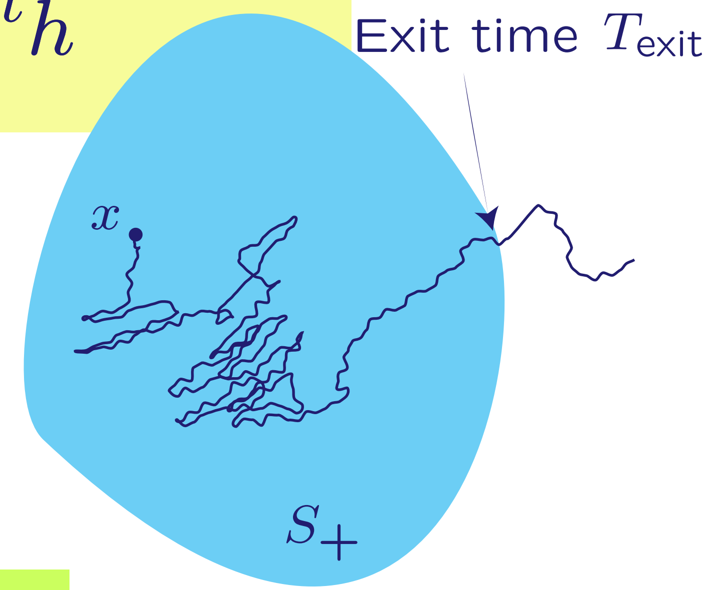
Under a density assumption a global LDP is established,
the rate function is identified as relative entropy,
and

the spectrum is discrete

Kontoyiannis & M. '05

Outline

Metastability & $P^t h = e^{-\Gamma t} h$



In continuous time,

$\Gamma > 0$, h is the second eigenfunction

$S_+ =$ connected component of $\{h > 0\}$

The exit time from a positive set S_+ is approximately exponentially distributed:

$$\mathsf{P}_x\Big\{\exp\Big(\beta(T_{\text{exit}} - T)\Big) \mid T_{\text{exit}} > T\Big\} = \frac{\Gamma}{\Gamma - \beta} \quad + \quad \mathcal{E}(x,T)$$

MGF for exponential rv

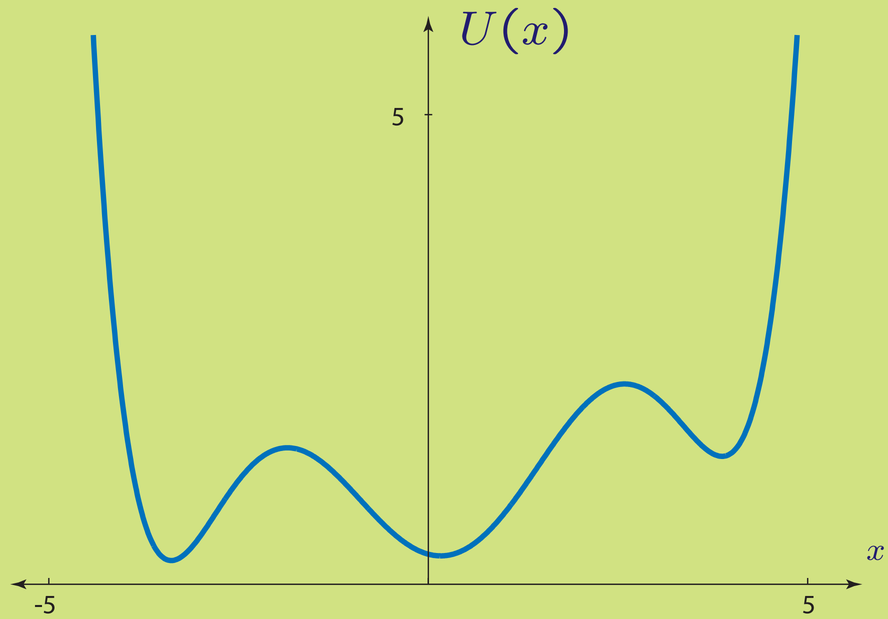
$$\mathsf{E}_x[f(\Phi(T)) \mid T_{\text{exit}} > T] = \check{\pi}(f) \quad + \quad \mathcal{E}(x,T)$$

Quasi steady-state on S_+

$$\mathcal{E}(x,T) = O\Big(v(x)h(x)^{-1}e^{-cT}\Big)$$

Huisinga M. Schuette '03

The three-well potential



The Smoluchowski equation

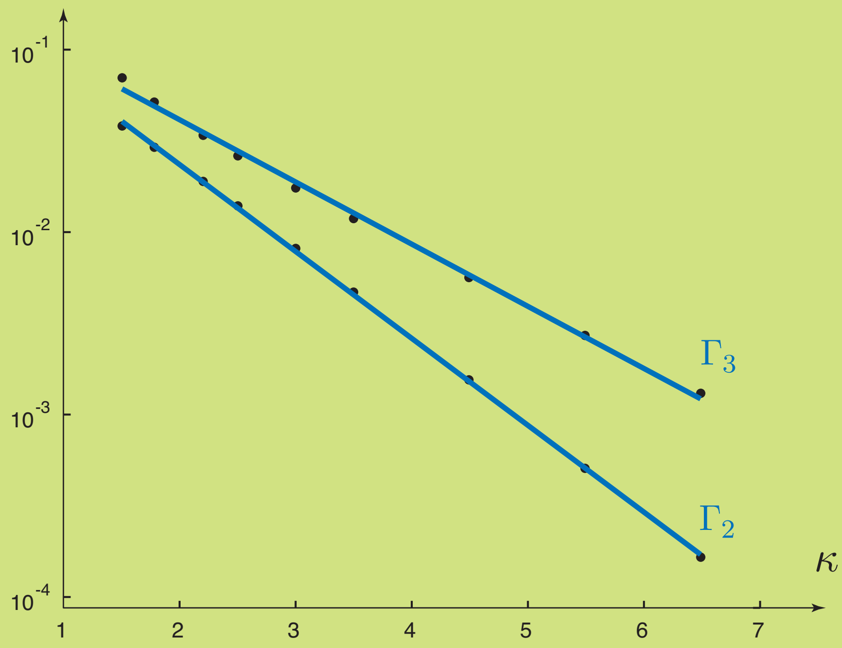
$$\mathrm{d}X = -\frac{1}{\gamma}\nabla U(X)\,\mathrm{d}t + \frac{\sigma}{\gamma}\,\mathrm{d}W$$

$V(x) = e^{\eta U(x)}$ solves (DV3)

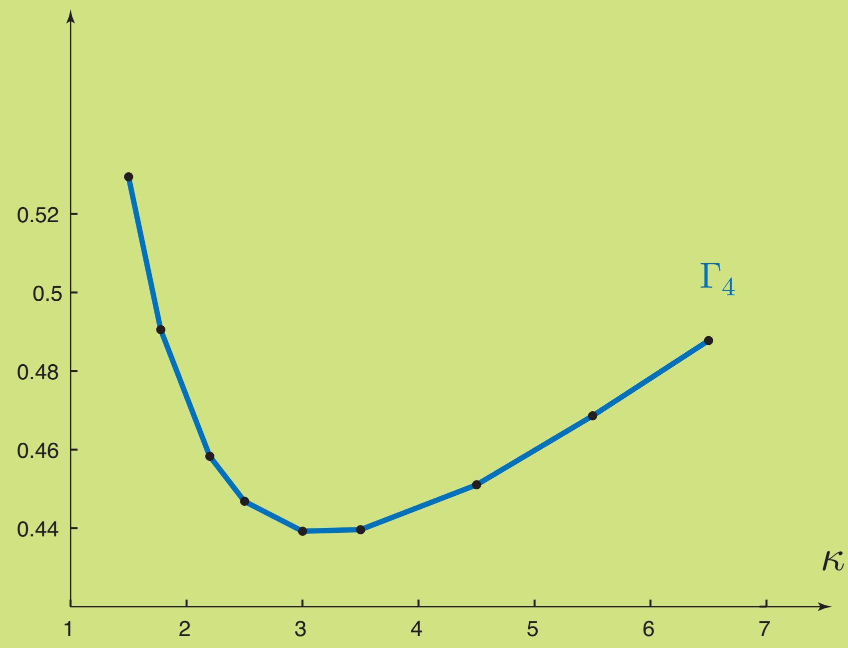
Invariant distribution with density $p(x) = \gamma e^{-\kappa U(x)}$

where $\kappa = 2\gamma/\sigma^2$ is the *inverse temperature*

Eigenvalues as a function of κ

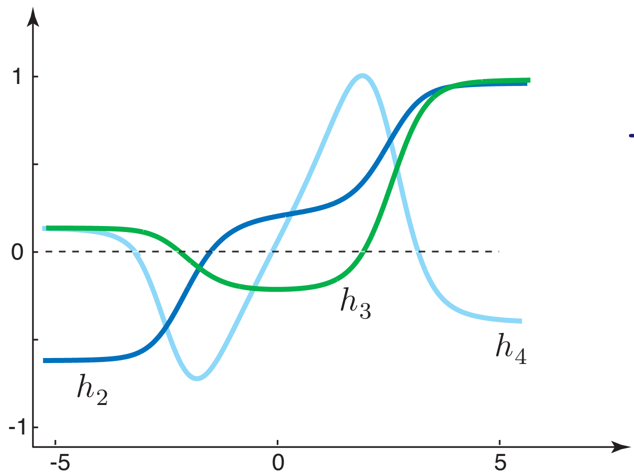


The second and third eigenvalues converge to zero

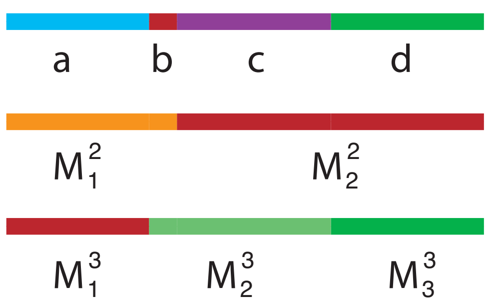


The fourth eigenvalue is bounded away from zero

Eigenfunctions and decompositions



Three eigenfunctions

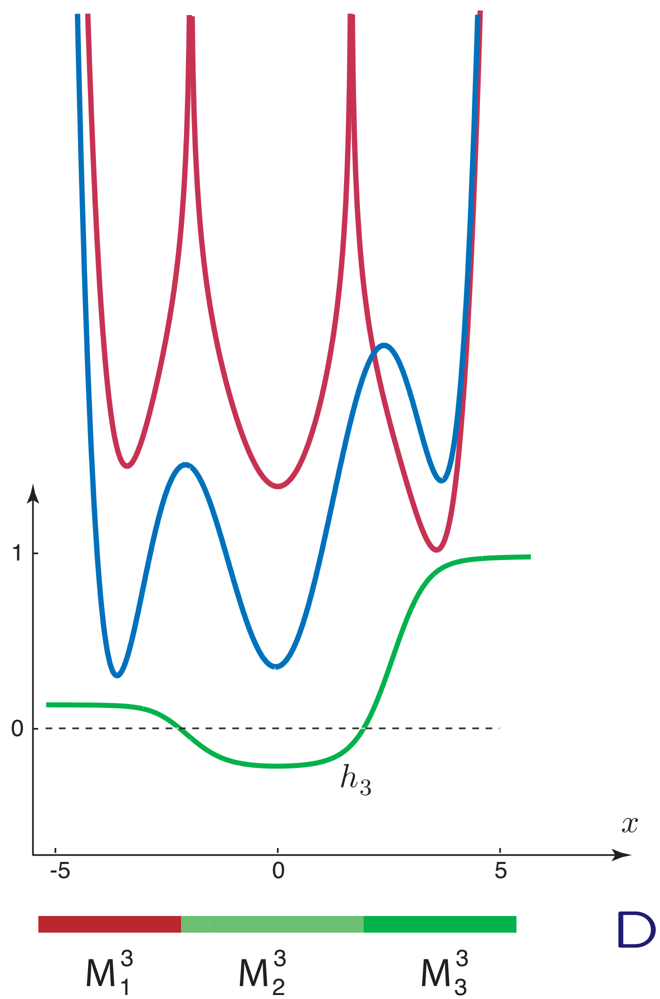


Decomposition based on zeros of $h_2 h_3$

Decomposition based on zeros of h_2

Decomposition based on zeros of h_3

The twisted potential



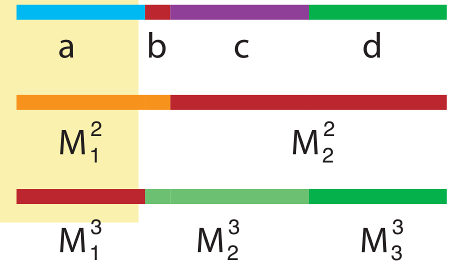
— The potential function $U(x)$

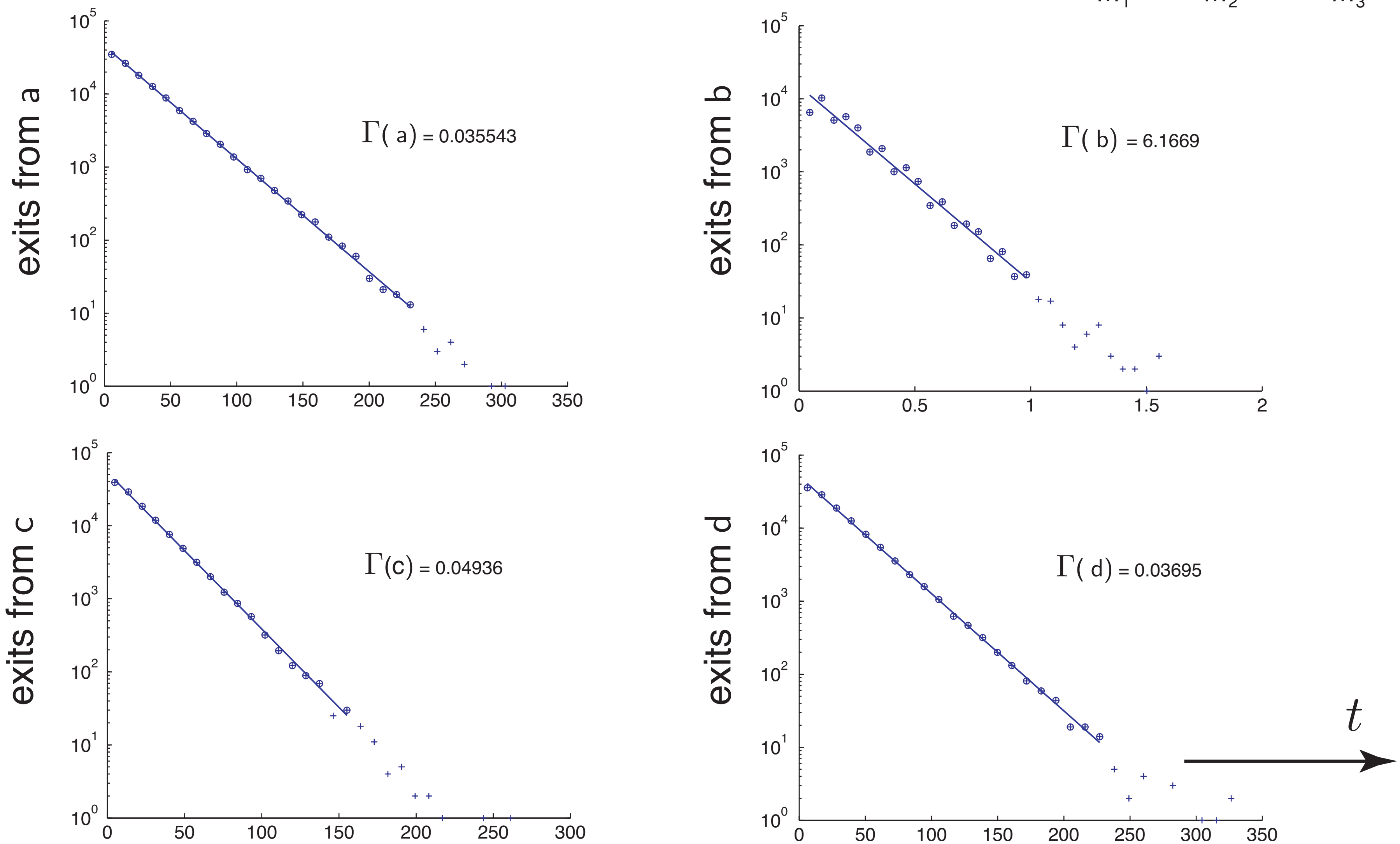
— The twisted potential function:

$$U_3(x) = U(x) - \sigma^2 \log(|h_3(x)|)$$

Decomposition based on zeros of h_3

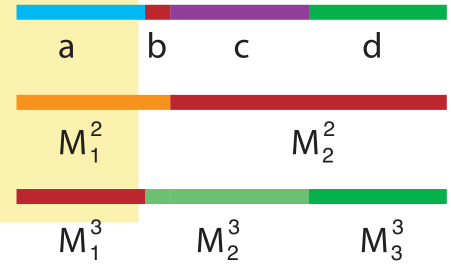
Exit time distributions

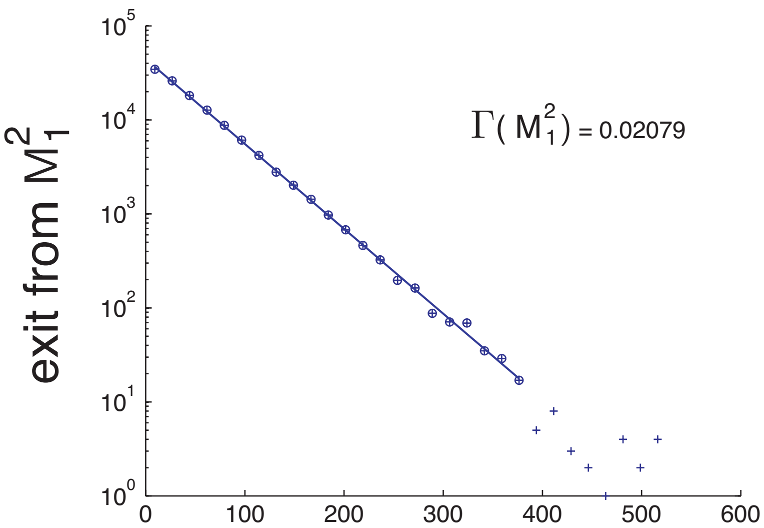


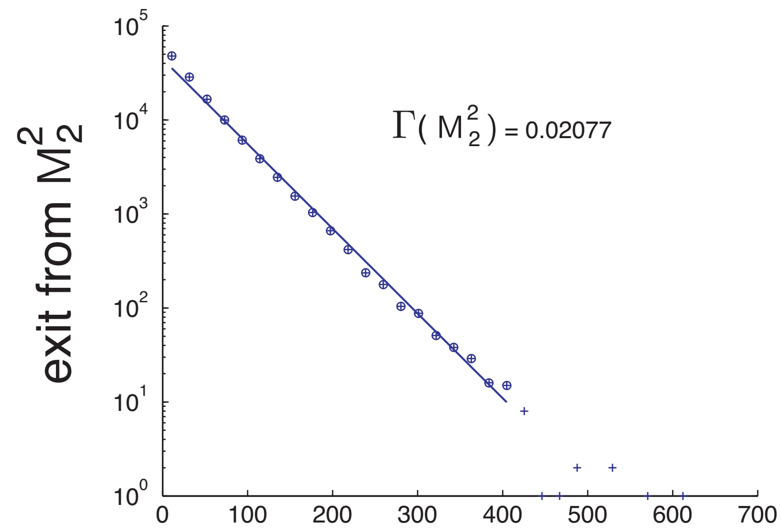


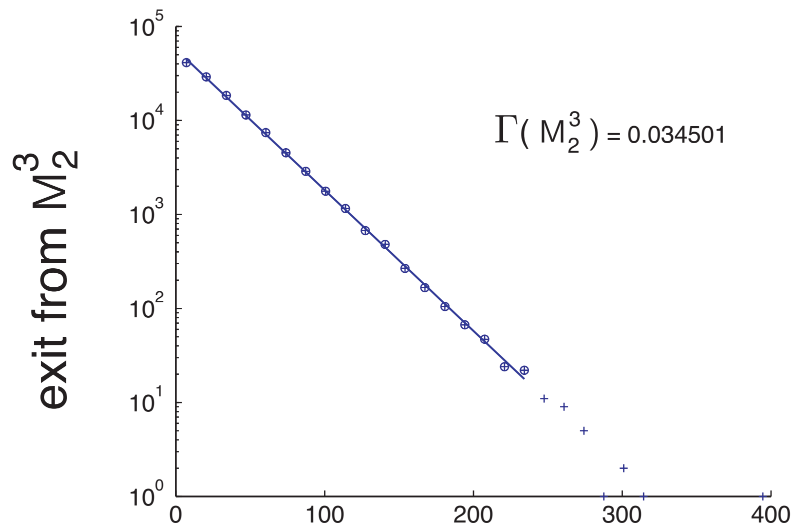
Exit time distribution for the sets $S = \{a, b, c, d\}$ and their combinations

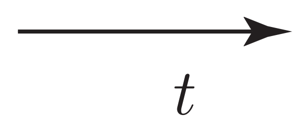
Exit time distributions











Outline

Finite approximations for P

(discrete time, densities, for simplicity)

Suppose that P has a continuous density,

$$P(x, dy) = p(x, y)\psi(dy)\,, \qquad x, y \in \mathsf{X}$$

On compact sets we can approximate,

$$p(x, y) \approx \sum_{i=1}^{n} s_i(x) p_i(y)\,,$$

$$Q(x, dy) = \sum_{i=1}^{n} s_i(x) \nu_i(dy)\,, \qquad x, y \in K.$$

Do we have $|||P - Q|||_v < \epsilon$ *?*

Can we assume that Q *is a transition kernel?*

Finite approximations for P

(discrete time, densities, for simplicity)

Suppose that $V, W : \mathsf{X} \to [1, \infty)$ are continuous with compact sublevel sets.

Suppose there exists $b < \infty$, satisfying (DV3):

$$\mathcal{H}(V) \triangleq \log(Pe^V) - V \leq -W + b$$

Then, for any $\epsilon > 0$, there exists $r \geq 1$ and a finite approximation on

$$K = \{x : W(x) \leq r\}.$$

Yes! We CAN assume that Q is a transition kernel

we have

$$|||P - Q|||_v < \epsilon$$

$$\sup_{t \geq 1} |||P^t - Q^t|||_v = O(\epsilon)$$

Dynamics for Q

$$Q(x, dy) = \sum_{i=1}^{n} s_i(x)\nu_i(dy)$$

A realization of the Markov chain Ψ under Q :

Process is a randomized function of a chain I on $\{1, \dots, n\}$

$$\Psi(0) = x$$

$$s(x) = [s_1(x), \dots, s_n(x)]^T$$

$$\mathsf{P}\{I(0) = i\} = s_i(x)$$

$$\mathsf{P}\{\Psi(1) \in A \mid I(0) = i\} = \nu_i\{A\}, \qquad A \in \mathcal{B}, \quad i = 1, \dots, n$$

$$\mathsf{P}\{I(1) = j \mid \Psi(0),\ \Psi(1)\} = s_j(\Psi(1))$$

Dynamics for Q

$$Q(x,dy) = \sum_{i=1}^{n} s_i(x)\nu_i(dy)$$

The process $I(k)$ is a Markov chain on $\{1,\dots,n\}$

Its transition law:

$$\mathsf{P}\{I(k) = j \mid I(k-1) = i\} = \nu_i(s_j)\,, \qquad i,j \in \{1,\dots,n\}$$

The finite chain serves as a sufficient statistic

$$\mathsf{P}\{\Psi(k) \in A \mid I(k-1) = i, \Psi(0),\dots,\Psi(k-1)\} = \nu_i(A)$$

Conclusion:

Anything that is true for I is also true for Ψ

and, with a bit more effort, true for Φ

Spectral theory for Q

Q has $\leq n$ eigenfunctions $\{h_1, \ldots, h_n\} \subset$ *span* $\{s_1, \ldots, s_n\}$

$\leq n$ eigenmeasures $\{\mu_1, \ldots, \mu_n\} \subset$ *span* $\{\nu_1, \ldots, \nu_n\}$

P has a finite spectrum within the region

$$B_\epsilon = \{z \in \mathbb{C} : |z| \geq \epsilon\},$$

and eigenvectors are approximated by those of Q

In particular,

$$P^t \rightarrow \mathbf{1} \otimes \pi$$
$$Q^t \rightarrow \mathbf{1} \otimes \varpi$$

with $\|\pi - \varpi\|_v = O(\epsilon)$.

Outline

Conclusions

- ▷ The DV drift condition allows a finite approximation of P *provided one considers the weighted* L_∞*-norm*
- ▷ Strong conclusions are possible without reversibility. Complex eigenvalues to be explored.
- ▷ Far weaker assumptions are possible, such as a relaxation of the one-step density conditions
- ▷ Recent applications to products of random matrices, and spectral decompositions for diffusions
- ▷ It seems likely that the finite approximations will find application in other areas **!**

References

I. Kontoyiannis and S.P. Meyn, Spectral Theory and Limit Theory for Geometrically Ergodic Markov Processes, Annals of Applied Prob., Volume 13, pp. 304-362, 2003.

J. Huang, I. Kontoyiannis, and S.P. Meyn, The O.D.E. Method and Spectral Theory of Markov Operators. Proceedings of the Second Kansas Workshop on Stochastic Theory - Adaptive Control, 2001.

W. Huisinga, S. P. Meyn, and C. Schuette, Phase Transitions & Metastability in Markovian and Molecular Systems. Annals of Applied Prob., Vol. 14, pp. 419-458, 2004

I. Kontoyiannis and S.P. Meyn, Large Deviations Asymptotics and the Spectral Theory of Multiplicatively Regular Markov Processes, Electron. J. Probab., Vol. 10, pp. 61--123 (electronic), 2005.